

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2004, 16:17:09 ; Search time 3726 Seconds
(without alignments)
6718.759 Million cell updates/sec

Title: US-09-976-673-11
Perfect score: 687
Sequence: 1 atggtgagcggcctgctgaa.....tgcccgagaaggccaactga 687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_gss1:*
8: gb_gss2:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|------------|-------|-------------|--------|----|-------------|---------------------|--|
| Result No. | Score | Query Match | Length | ID | Description | | |
| 1 | 99.4 | 14.5 | 550 | 5 | BQ551001 | BQ551001 H4006G01- | |
| 2 | 71 | 10.3 | 645 | 1 | AI389106 | AI389106 GH20192.5 | |
| 3 | 68.6 | 10.0 | 869 | 7 | CK159167 | CK159167 FGAS04056 | |
| 4 | 67.6 | 9.8 | 639 | 1 | AJ615477 | AJ615477 AJ615477 | |
| 5 | 67.4 | 9.8 | 675 | 4 | BI382638 | BI382638 BFLG2_000 | |
| 6 | 66.8 | 9.7 | 944 | 9 | CC649516 | CC649516 OGLCL16TH | |
| 7 | 66.4 | 9.7 | 700 | 7 | CF864358 | CF864358 pszs009xp | |
| 8 | 65.8 | 9.6 | 624 | 8 | BZ422085 | BZ422085 id49d07.b | |
| 9 | 65.8 | 9.6 | 982 | 6 | CC424389 | CC424389 PUHMK53TB | |
| 10 | 65.2 | 9.5 | 794 | 6 | CB651092 | CB651092 OSJNEB150 | |
| 11 | 64.8 | 9.4 | 697 | 5 | BU626888 | BU626888 UI-H-FT0- | |
| 12 | 64.6 | 9.4 | 821 | 7 | CK199961 | CK199961 FGAS00846 | |
| 13 | 64.2 | 9.3 | 613 | 4 | BI386976 | BI386976 BFL26_001 | |
| 14 | 64.2 | 9.3 | 643 | 4 | BI386699 | BI386699 BFL26_001 | |
| 15 | 64 | 9.3 | 681 | 9 | CNS02EOD | AL193990 Tetraodon | |
| 16 | 64 | 9.3 | 688 | 9 | CG038305 | CG038305 PUFNQ30TD | |
| 17 | 63.8 | 9.3 | 411 | 2 | BF201975 | BF201975 WHE1759-1 | |
| 18 | 63.8 | 9.3 | 600 | 2 | BE417750 | BE417750 MUG024.B0 | |
| 19 | 63.8 | 9.3 | 772 | 5 | BQ752733 | BQ752733 WHE4118 E | |
| 20 | 63.6 | 9.3 | 872 | 8 | CC413363 | CC413363 PUHPU48TD | |
| 21 | 63.6 | 9.3 | 956 | 9 | CG365299 | CG365299 OGWF62TV | |
| 22 | 63.4 | 9.2 | 925 | 9 | CNS0091P | AL053013 Drosophila | |
| 23 | 62.8 | 9.1 | 1006 | 9 | CC719047 | CC719047 OGULN81TV | |
| 24 | 62.6 | 9.1 | 599 | 1 | AI832082 | AI832082 tdl2c04.x | |

| | | | | | | |
|----|------|-----|------|---|----------|--------------------|
| 25 | 62.6 | 9.1 | 689 | 4 | BI380055 | BI380055 BFLG1_001 |
| 26 | 62.4 | 9.1 | 1036 | 7 | CK162903 | CK162903 FGAS01551 |
| 27 | 62.2 | 9.1 | 638 | 4 | BJ269933 | BJ269933 BJ269933 |
| 28 | 62.2 | 9.1 | 742 | 6 | CB649099 | CB649099 OSJNEB12N |
| 29 | 62.2 | 9.1 | 780 | 7 | CF554490 | CF554490 Ggw74 Inf |
| 30 | 62 | 9.0 | 786 | 6 | CB646357 | CB646357 OSJNEB08M |
| 31 | 61.8 | 9.0 | 557 | 4 | BJ233355 | BJ233355 BJ233355 |
| 32 | 61.6 | 9.0 | 590 | 6 | CA030137 | CA030137 HX06D04r |
| 33 | 61.4 | 8.9 | 540 | 4 | BM135283 | BM135283 WHE0498.B |
| 34 | 61.4 | 8.9 | 635 | 6 | CA254345 | CA254345 SCBFLL411 |
| 35 | 61.2 | 8.9 | 724 | 7 | CF871694 | CF871694 trico28xf |
| 36 | 61.2 | 8.9 | 739 | 6 | CB645917 | CB645917 OSJNEB07P |
| 37 | 61.2 | 8.9 | 777 | 6 | CB901927 | CB901927 trico28xf |
| 38 | 61 | 8.9 | 571 | 4 | BI387406 | BI387406 BFL26_002 |
| 39 | 61 | 8.9 | 629 | 1 | AB118745 | AB118745 AB118745 |
| 40 | 61 | 8.9 | 661 | 4 | BI385436 | BI385436 BFL26_000 |
| 41 | 61 | 8.9 | 687 | 4 | BI387917 | BI387917 BFL26_002 |
| 42 | 61 | 8.9 | 813 | 6 | CB657967 | CB657967 OSJNEG13K |
| 43 | 60.8 | 8.9 | 502 | 4 | BI135401 | BI135401 6HRm115.6 |
| 44 | 60.6 | 8.8 | 559 | 7 | CF847678 | CF847678 psMA002xB |
| 45 | 60.6 | 8.8 | 759 | 8 | BZ577913 | BZ577913 msh2_5633 |

ALIGNMENTS

RESULT 1
BQ551001
LOCUS
DEFINITION
H4006G01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
H4006G01 5', mRNA sequence.
ACCESSION
BQ551001
VERSION
BQ551001.1 GI:21451887
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 (bases 1 to 550)
AUTHORS
VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T.,
Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
CDNA clone set
JOURNAL
Genome Res. 12 (12), 1999-2003 (2002)
MEDLINE
22354164
PubMed
12466305
COMMENT
Other_ESTs: H4006G01-3
Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
Plate: H4006 row: G column: 01
Seq primer: -21M13 Reverse
High quality sequence stop: 550
POLYA=NO.

FEATURES
source
1..550
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4006G01-5"
/db_xref="taxon:10090"
/clone="H4006G01"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more

than 20 cdna libraries."

```
ORIGIN
Query Match      14.5%; Score 99.4; DB 5; Length 550;
Best Local Similarity 55.3%; Pred. No. 2e-10;
Matches 215; Conservative 0; Mismatches 171; Indels 3; Gaps 1;

QY 13 CTGCTGAAGGAGATGTCGGCATCAAGATGATGATGAGGGCACCCTGTAACGGCCACTAC 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 CTGTTACCGGGTGTGTCCTCATCTCTGTCGAGCTGGACGGCGACGTAAACGGCCACAAG 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 73 TTCAAGTGGAGGGGAGGGCGACGGCAACCCCTTCGCCCGGACCCACGAGCATGAGAATC 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTC 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 133 CACGTGACCGAGGGC---GCCCCCTGCGCTTCGCCCTTCGACATCCTGGCCCCCTGCTGC 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 ATCTGCACCAACGGGAAGCTGCCGTGCCCTGGCCCCACCTCTGTCGACCACTGACCTAC 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 190 GAGTACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCGGACTTCTTCAAGCAG 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 GCGGTGCAGTCTTACGCCGCTACCCCGACCACTGAAGCAGCAGCACTTCTTCAAGTCC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 250 AGCTTCCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGACGGCGCATCTCG 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 GCCATGCCCCGAGGCTACGTCAGGAGCGGCACCATCTTCTTCAAGGACGACGGCAACTAC 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 310 ACCGCCACACGAGACACAGCCTGGAGGGCAACTGCTGATCTTCAAGTGAAGGTGCAC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 AAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAG 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 370 GGCACCAAACTTCCCGCGACGGCCCGCT 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 GGCATCGACTTCAAGGAGGACGGCAACAT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 2
LOCUS AI389106 645 bp mRNA linear EST 23-APR-2001
DEFINITION GH20192.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH20192 5, mRNA sequence.
ACCESSION AI389106
VERSION AI389106.2 GI:13758704
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 645)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT On Jan 28, 1999 this sequence version replaced gi:4203117.
Other_ESTs: GH20192.3prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: GH.201 row: H column: 8
High quality sequence stop: 627
POLYA=No.
Location/Qualifiers
1..645
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH20192"
/sex="male and female"
/dev_stage="adult"
```

FEATURES

Source

```
/lab_host="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pOT2"
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."
```

```
ORIGIN
Query Match      10.3%; Score 71; DB 1; Length 645;
Best Local Similarity 48.6%; Pred. No. 0.00015;
Matches 194; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 62 ACGGCCACTACTTCAAGTCCGAGGGCGGAGGGGACGGCAACCCCTTCGCCGGCACCCAGA 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 ACGACCACTCCACCATGTGCTAGGAAACCCAGACCAACACGCTGTCCCTACTACCACA 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 122 GCATGAGAATCCACGTGACCGAGGGCGGCCCTTCGCCCTTCGACATCCTGGCCC 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 ACAACCACAACGTTTCCCCAACTACCACTACTACCACTACCACTACCACTACCACTACCA 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 182 CCTGCTGCGAGTACGCGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCGACTTCT 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 ACCACCACAACCTACGACAACAACAACCTACCACTACCACTACCACTACCACTACCACTAC 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 242 TCAAGCAGAGCTTCCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCG 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 CCGACAACAACCCACCAACAACAACAACGACGACGACGACGACGACGACGACGACGACGAC 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 302 GCATCCTGACCGCCCAACAGGACACCACTTCGTCGACCAACCGCCGAGATCCCGACTTCT 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 ACAACCACGACGACGACAACAACAACCACTACCACTACCACTACCACTACCACTACCACTAC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 362 AGGTGCACGGCACCAACTTCCCCCGCGACGGCCCGCTGTATGAAGAACAGAGCGCGGCT 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 ACCACCACCAACAACCTACTACTACGACACCACTACGACAACAACAACCACTACCACTAC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 422 GGGAGCCCGACGACCGAGGTGTGTACCCCGAGAACGGCG 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 560 ACAACAACCAACGACGACAACAACAACAACCACTACCACTACCACTACCACTACCACTAC 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

CKI59167/c

LOCUS

DEFINITION

CKI59167

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Triticum aestivum (bread wheat)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.

REFERENCE

AUTHORS

1 (bases 1 to 869)

Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D., Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A., Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D., Penniket,C., Roach,J.L. and Sarhan,F.

Functional Genomics of Abiotic Stress In Wheat and Canola Crops

Unpublished (2003)

Contact: Wm L Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science

1C101 Engineering Building, 57 Campus Drive, Saskatoon,

Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas_est@cs.usask.ca

This sequence is the direct result of the Base calling software

Phred (default parameters). It is the raw base calls. To aid in the

identification of the high quality insert the software Lucy

(default parameters) has been run on this sequence. Lucy identified

the region [128,636].

JOURNAL
MEDLINE
PUBMED
COMMENT

genomes
Genome Res. 13 (6A), 1056-1066 (2003)
22683279
12799346
Contact: Panopoulou G
laboratory 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: panopoul@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones giving the same hybridisation pattern with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well the coordinates of the
rest of the clones assigned to the same fingerprint cluster as the
clone from which the above EST is generated is available at the
amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
Clones and filters are distributed via the Resource Center/Primary
Database of the German Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCAGGCTTTACACTTTATGCTTCGGCTCG 3' (M13RSP)
BACKWARD: 5' GCTATTACCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)
Insert Length: 1 Std Error: 0.00
Seq primer: 5'-CCGGTCCGGAATCCCGGT-3' pSport3/86
High quality sequence stop: 675.
Location/Qualifiers
1. .675
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MPMGp498F0258"
/tissue_type="whole embryo"
/dev_stage="5-6 hrs (gastrula stage)"
/lab_host="E.coli, XL1 blue"
/clone_lib="Amphioxus 5-6 hrs cDNA library (Name
convention: BFLG or MPMGP498)"
/note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5');
Site 2: NotI, BamHI, HindIII (3'); OligodT primed and
directionally cloned in pSport1 vector using a NotI
(5')-pGACTAGTTCTAGATCGGAGCGCGCCCC (T)15-3' and a SalI 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

FEATURES
source

ORIGIN

Query Match 9.8%; Score 67.4; DB 4; Length 675;
Best Local Similarity 48.1%; Pred. No. 0.00084;
Matches 191; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

Qy 36 CAAGATGTACATGGAGGGACCGTGAACGGCCACTTCAAGTGGAGGGCGAGGGCGA 95
Db 106 CGACATCCACCTTACCGGCTCCATCAACGGCCACGAGTTCGACATGGTGGGGGAGGAAA 165

Qy 96 CGGCAACCCCTTCGCGGACCCAGAGCATGAGATCCACGTGACCGAGGGCGCCCCCT 155
Db 166 AGGCGACCCGAACCGCGCTCGCTGGTGACCAAGCAATCCACCAAGGTTCCCTGAA 225

Qy 156 GCCCTTCGCTTCGACATCCCTGGCCCCCTGCTGCGAGTACGGCAGAGGACCTTCGTGCA 215
Db 226 GTTCTCCCCCTACTTGTATGATCCCCCACCCTCGGGTACGGGTACTACCACTCCCTA 285

Qy 216 CCACACCGCCGAGATCCCGGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGA 275
Db 286 CCGGACGACCCCTCGCCCTTCCAGGTCTCCATGTTGGAAGGATCGGGGTATGCAGTCTA 345

Qy 276 GAGAACACACCTACGAGGACGGCGGCATCCTGACCGCCACCAGGACACAGCCTGGA 335
Db 346 CCGCGTGTTCGACTTTGAAGACGGAGGCAAGCTGACTACCGAGTTTAAGTACTCCTACGA 405

Qy 336 GGGCAACTGCCTGATCTACAGGTGAAGGTGCACGGCACCAACTTCCCCCGGACGGCCC 395

Db 406 GGGTTCCCATATCAAGGCCGACATGAAGCTGATGGGAAGCGGTTTCCCTGACGACGGGCC 465

Qy 396 CGTGATGAAGAACAGAGCGGGGCTGGAGCCACG 432
Db 466 AGTCATGACCAGCCAGATTGTCGACCAGGACGGCTGC 502

RESULT 6
CC649516
LOCUS
DEFINITION
OGLCL16TH ZM_0.7.1.5_KB Zea mays genomic clone ZMMBma0322D08,
genomic survey sequence.
ACCESSION
CC649516
VERSION
CC649516.1 GI:32052184
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 944)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLCL16TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: Sheared ends.
Location/Qualifiers
1. .944
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0322D08"
/clone_lib="ZM_0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
1. .944
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0322D08"
/clone_lib="ZM_0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.7%; Score 66.8; DB 9; Length 944;
Best Local Similarity 45.7%; Pred. No. 0.0011;
Matches 273; Conservative 0; Mismatches 322; Indels 3; Gaps 1;

Qy 30 GCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTACTTCAAGTGCAGGGCGA 89
Db 129 GCGCTCGCCCTGTTCTACGCGCGCGCGCTCGTCCGACCCCTCGGCGGCAC 188

Qy 90 GGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGATCCACGTGACCGAGGGCGC 149
Db 189 CGGCTCGCCAAACGCCGACTCGTCTACTTCAACGGCGCGCTCCTCGCCATGTCCGAGGA 248

Qy 150 CCCCCTGCCCTTCGCCCTTCGACATCCTGGCCCCCTGCTGCGAGTACGGCAGGACCTT 209
Db 249 CGACTGCCGTACCAACGTCGCGTTCGCGGACGAGCGGCGACCTCGAGACCGTCCGCGCTA 308

Qy 210 CGTGACACACACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCAC 269
Db 309 CGACTTCGACGGCCAGCTCGGCTGCCCATGATCGCGCACCCCAAGCTGACCGCGCCAC 368

Qy 270 CTGGAGAGAACCAACCACTACGAGGACGGCGGCATCTCTGACCGCCACAGGACACAG 329
Db 369 CGGGAGCTGCACCGCGCTCAGTACGAGGTCTGAGGAGGCCCTACCTCAAGTACTTCTA 428

Qy 330 CCTGGAGGGCAACTGCCTGA--TCTACAAGTGAAGGTGCACGGCACCAACTTCCCCGC 386

Db 429 CTTACGGCCCGACCGCACCAAGTCCGACGACGTGGAGATCCCGCTGGCCCGACCCACCAT 488

QY 387 CGACGGCCCCGTGATGAAGAACAAAGACGGCGGCTGGAGCCGACGACCCGAGGTGGTGA 446

Db 489 GATCCACGACTTCGCCATCACCGAGAACTTGGTCGTGGTGCCTCCGACACCCAGGTGGTGT 548

QY 447 CCCCAGAACCGCGTGTGTGCGGCGGAAACGTGATGGCCCTGAAGGTGGCGACCGGCA 506

Db 549 CAAGCTGCAGAGATGCTGCGCGCGGTCGCCCGTGGTGTGACAGGGAGAACGTC 608

QY 507 CCTGATCTGCCACCACTACACGACTACCGGCTCCAGATGCTGCGGAAGAAGACGAG 566

Db 609 GCGCTTCGGGCTCCTCCGAAAGCGCGCGGACGCGTCGGAGATGGCGTGGTGGACGT 668

QY 567 GCCCGGTTTCCACTTCACCGACATCCGGCTCCAGATGCTGCGGAAGAAGACGAG 624

Db 669 GCCGGAAGTCTTCTGCTTCCACCTGTGGAACGCGTGGAGGACGAGGCGACGGCGGAG 726

RESULT 7

CF864358

LOCUS

DEFINITION

psZS009xP17f USDA-IFAFS:Expression of Phytophthora sojae genes during infection and propagation_szs Phytophthora sojae cDNA clone szS009P17 5, mRNA sequence.

ACCESSION

CF864358

VERSION

CF864358.1 GI:38118984

KEYWORDS

EST.

SOURCE

Phytophthora sojae

ORGANISM

Phytophthora sojae

REFERENCE

1 (bases 1 to 700)

AUTHORS

Tyler,B.

TITLE

Tyler,B. Not Published

JOURNAL

Unpublished (2003)

COMMENT

Contact: Tyler B

Tyler lab

VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmt Tyler@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 009 row: P column: 17

Seq primer: BK reverse primer

High quality sequence stop: 700.

FEATURES

source

1. 700

/organism="Phytophthora sojae"

/mol_type="mRNA"

/db_xref="taxon:67593"

/clone="szS009P17"

/tissue_type="Zoospores"

/cell_line="P6497"

/dev_stage="Free swimming"

/clone_lib="USDA-IFAFS:Expression of Phytophthora sojae genes during infection and propagation_szs"

/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

Best Local Similarity 45.7%; Pred. No. 0.0014;

Matches 269; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 66 CCACTACTTCAAGTCCGAGGGCGGAGCGGACGGCAACCCCTTCGCCGGCACCCAGACAT 125

Db 103 CCGCCACACCAAGAGCCACAGCGTGGATGGCATCAGCCCGCGCCAGCCACTCCACCGT 162

QY 126 GAGATCCACGTGACCGAGGGCGCCCCCTGCCCTTCGACATCCTGGCCCCCTG 185

Db 163 GAG---CTCCGTCGCCACGCTCGCCGCGCCGACCAAGACGATCCGCCACCGCCGCG 219

QY 186 CTGCGAGTACGGCAGCAGGACCTTCGTGCACCAACACCGCCGAGATCCCGACTTCTTCAA 245

Db 220 CCTCTCCACGGGTCAAATGGCTGACGAGGTGTCCTCCCAAGGAGCGGATCGACGTCTA 279

QY 246 GCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGACGGCGCAT 305

Db 280 CGACTACTCATGAGCGCATCTCGAGTCGCCCGCTCATCGGCGTGTGTGCGGCTCCGG 339

QY 306 CCTGACCGCCCAACGAGACACCAAGCCCTGGAGGGCAACTGCCTGATCTACAAGGTGAAGT 365

Db 340 CCTCGAGGACTAGCAACTGCCTGTCTCAACACGAGACCATCAAGTACGAGGACATCCC 399

QY 366 GCACGGCACCAACTTCCCGCCGACGCGCCCGCTGATGAAGAACAAAGACGCGGCTGGGA 425

Db 400 GCAGTTCCCGGCTCCACCGTCGAGGGCCACGCGGAGAGTCTGCTTCGGCGACCTGGA 459

QY 426 GCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGATGGC 485

Db 460 GGGCTTCCGCTGCTGTCATGCGCGCGGCTTCCACTGTCTACGAGGGCTACGCCATGCG 519

QY 486 CCTGAAGGTGGGCGACCGGCACCTGATCTGCCACCACTACACCAAGTACCGGAGCAAGAA 545

Db 520 CGAGACGGCGTGCCTCATTCGCGTTCATGATCTCCTCGGCATCAAGTACTTGTGTGAC 579

QY 546 GGCGGTGCGCGGCTGACCATGCCCGCTTCCACTTCCACCGACATCCGGCTCCAGATGCT 605

Db 580 CAACGCGCGGCGGCTGAAACCGGACTTCAACGTGGCGGACGTCATGATCCTGAACGA 639

QY 606 GCGGAAGAAAGAGACGAGTACTTTCGAGTGTACGAGGCGGCGGCTCCAGTGGC 653

Db 640 CCACCTGAACGTGCGCGGCTGTGCGGTTCAGCACCCGCTCATCGGGCC 687

RESULT 8

BZ422085

LOCUS

DEFINITION

id49d07.b1 WGS-SbicolorF (DH5a methyl filtered) Sorghum bicolor genomic clone id49d07 5', genomic survey sequence.

ACCESSION

BZ422085

VERSION

BZ422085.1 GI:26370029

KEYWORDS

GSS.

SOURCE

Sorghum bicolor (sorghum)

ORGANISM

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

1 (bases 1 to 624)

AUTHORS

Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.

TITLE

Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)

JOURNAL

Unpublished (2002)

COMMENT

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: id49 row: d column: 07

Seq primer: -21M13UnivFwd

Class: shotgun

High quality sequence stop: 624.

FEATURES

Location/Qualifiers

1. 624

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/db_xref="taxon:4558"

/clone="id49d07"

/lab_host="DH5a"

/clone_lib="WGS-SbicolorF (DH5a methyl filtered)"

/note="Site_1: Xba I; Site_2: Xba I; The vector was

/tissue_type="whole embryo"
/dev_stage="26 hrs (neurula stage)"
/lab_host="E.coli, Xll blue"
/clone_lib="Amphioxus 26hr cDNA library (Name convention:
BFL26 or MPMGp531)"
/note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI,
EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT
primed and directionally cloned in pSport1 vector using a
NotI (5'-pGACTAGTTCTAGATCGGAGCGGCC (T)15-3' and a
SalI 5'-TCGACCCACGGCTCCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 9.3%; Score 64.2; DB 4; Length 613;
Best Local Similarity 47.6%; Pred.No. 0.0039;
Matches 189; Conservative 0; Mismatches 208; Indels 0; Gaps 0;
QY 36 CAAGATGATACATGAGGGCACCCTGACCGCCACTACTTCAAGTCCGAGGGCGAGGCGGA 95
Db 88 CGACATCCACCTTACGGCTCCATCAACGGCCACGAGTTCGACATGGTGGGGGAGGAAA 147
QY 96 CGGCAACCCCTTCGCCGGCACCCAGAGCATGAGATCCACGTGACCGAGGGCGGCCCT 155
Db 148 AGCGACCCGACCGCGCTCGTGGTGACACAGCGGAATCCACCAAGGTGCCCTGAA 207
QY 156 GCCCTTCGCTTCGACATCTCTGGCCCCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCA 215
Db 208 GTTCTCTCCTACTTGTATGATCCCCACCTCGGTACGGGTACTACAGTACCTCCCTA 267
QY 216 CCACACCGCGGAGATCCCGACCTTCTCAAGCAGAGAGTTCCTCCCGAGGGCTTACCTGGGA 275
Db 268 CCGGACCGACCTCGCTTCCAGGCTCCATGTTGGAAGGATCGGGTATGACTCTA 327
QY 276 GAGAACACCACTACGAGGACCGCGGCATCTGACCGCCACCGACAGGACACCGCTGGA 335
Db 328 CCGCGTGTCTGACTTTGAAGACGAGGCAAGCTGACTACCGAGTTTAAGTACTCCTACGA 387
QY 336 GGGCAACTGCTGATCTACAAGGTGAAGGTGCACGGCACCAACTTCCCGCGAGCGGCC 395
Db 388 GGGTCCCATATCAAGGCCGACATGAAGCTGATGGGAGCGGTTTCCCTGATGACGGCCC 447
QY 396 CGTGATGAAGAACAGAGCGCGGCTGGGAGCCCGCAGC 432
Db 448 AGTCATGACGACCGAGATTGTGACCGACGAGCGGCTGC 484

RESULT 14
BI386699
LOCUS
DEFINITION
BFL26_001402 Amphioxus 26hr cDNA library (Name convention: BFL26 or
MPMGp531) Branchiostoma floridae cDNA clone MPMGp531L1370 5', mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
BI386699 643 bp mRNA linear EST 26-AUG-2003
BFL26_001402 Amphioxus 26hr cDNA library (Name convention: BFL26 or
MPMGp531) Branchiostoma floridae cDNA clone MPMGp531L1370 5', mRNA
sequence.
BI386699 GI:30921604
BI386699.1
EST.
Branchiostoma floridae (Florida lancelet)
Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
1 (bases 1 to 643)
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A.J.,
Herwig, R., Vingron, M. and Lehrach, H.
New evidence for genome-wide duplications at the origin of
vertebrates using an amphioxus gene set and completed animal
genomes
Genome Res. 13 (6A), 1056-1066 (2003)
22683279
12799346
Contact: Panopoulou G
Laboratory 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones giving the same hybridisation pattern with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well the coordinates of the
rest of the clones assigned to the same fingerprint cluster as the
clone from which the above EST is generated is available at the
amphioxus project site at: http://www.molgen.mpg.de/amphioxus/
Clones and filters are distributed via the Resource Center/Primary
Database of the German Genome Project (http://www.rzpd.de)
PCR Primers
FORWARD: 5' CCCAGGCTTTACACTTATGTTCCGGCTCG 3' (M13RSP)
BACKWARD: 5' GCTATTACGCCAGTGGCGAAAGGGGATGTG 3' (M13FSP)
Insert Length: 1 Std Error: 0.00
Seq primer: 5'-CCGGTCCGGAATTCGGGT-3' pSport3/86
High quality sequence stop: 643.

FEATURES

Location/Qualifiers
1..643
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MPMGp531L1370"
/tissue_type="whole embryo"
/dev_stage="26 hrs (neurula stage)"
/lab_host="E.coli, Xll blue"
/clone_lib="Amphioxus 26hr cDNA library (Name convention:
BFL26 or MPMGp531)"
/note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI,
EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT
primed and directionally cloned in pSport1 vector using a
NotI (5'-pGACTAGTTCTAGATCGGAGCGGCC (T)15-3' and a
SalI 5'-TCGACCCACGGCTCCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 9.3%; Score 64.2; DB 4; Length 643;
Best Local Similarity 47.6%; Pred.No. 0.0039;
Matches 189; Conservative 0; Mismatches 208; Indels 0; Gaps 0;
QY 36 CAAGATGATACATGAGGGCACCCTGACCGCCACTACTTCAAGTCCGAGGGCGAGGCGGA 95
Db 77 CGACATCCACCTTACGGCTCCATCAACGGCCACGAGTTCGACATGGTGGGGGAGGAAA 136
QY 96 CGGCAACCCCTTCGCCGGCACCCAGAGCATGAGATCCACGTGACCGAGGGCGGCCCT 155
Db 137 AGCGACCCGACCGCGCTCGTGGTGACACAGCGGAATCCACCAAGGTGCCCTGAA 196
QY 156 GCCCTTCGCTTCGACATCTCTGGCCCCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCA 215
Db 197 GTTCTCTCCTACTTGTATGATCCCGACCTCGGTACGGGTACTACAGTACCTCCCTA 256
QY 216 CCACACCGCGGAGATCCCGACCTTCTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGA 275
Db 257 CCGGACCGACCTCGCTTCCAGGCTCCATGTTGGAAGGATCGGGGTATGAGTCTA 316
QY 276 GAGAACCACTACGAGGACCGCGGCATCTGACCGCCACCGACAGGACACCGCTGGA 335
Db 317 CCGCGTGTCTGACTTTGAAGACGAGGCAAGCTGACTACCGAGTTTAAGTACTCCTACGA 376
QY 336 GGGCAACTGCTGATCTACAAGGTGAAGGTGCACGGCACCAACTTCCCGCGAGCGGCC 395
Db 377 GGGTCCCATATCAAGGCCGACATGAAGCTGATGGGAGCGGTTTCCCTGACGACGGCCC 436
QY 396 CGTGATGAAGAACAGAGCGCGGCTGGGAGCCCGCAGC 432
Db 437 AGTCATGACGACCGAGATTGTGACCGACGAGCGGCTGC 473

RESULT 15
CNS02EOD/c
LOCUS
CNS02EOD 681 bp DNA linear GSS 01-SEP-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2004, 16:17:09 ; Search time 100 Seconds
(without alignments)
4883.120 Million cell updates/sec

Title: US-09-976-673-11
Perfect score: 687
Sequence: 1 atgttgagcgcctgtctgaa.....tgcccgagaaggccaactga 687

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 297.4 | 43.3 | 723 | 4 | US-10-152-296-1 |
| 2 | 253.4 | 36.9 | 1482 | 4 | US-09-977-897-1 |
| 3 | 253.2 | 36.9 | 720 | 4 | US-09-839-650-1 |
| 4 | 176.6 | 25.7 | 678 | 3 | US-09-459-956-6 |
| 5 | 153.8 | 22.4 | 801 | 3 | US-09-459-956-7 |
| 6 | 145.4 | 21.2 | 699 | 3 | US-09-459-956-5 |
| 7 | 121.6 | 17.7 | 1104 | 3 | US-09-277-716-30 |
| 8 | 121.6 | 17.7 | 1104 | 4 | US-09-609-161B-30 |
| 9 | 121.6 | 17.7 | 1279 | 3 | US-09-277-716-31 |
| 10 | 121.6 | 17.7 | 1279 | 4 | US-09-609-161B-31 |
| 11 | 110.6 | 16.1 | 720 | 3 | US-09-172-063-14 |
| 12 | 110.6 | 16.1 | 720 | 4 | US-09-316-919-15 |
| 13 | 110.6 | 16.1 | 720 | 4 | US-09-602-641-14 |
| 14 | 110.6 | 16.1 | 720 | 4 | US-09-316-920A-15 |
| 15 | 109.8 | 16.0 | 690 | 3 | US-09-459-956-2 |
| 16 | 109 | 15.9 | 717 | 4 | US-09-513-783A-43 |
| 17 | 109 | 15.9 | 717 | 4 | US-09-430-656-43 |
| 18 | 109 | 15.9 | 720 | 3 | US-08-911-825-3 |
| 19 | 109 | 15.9 | 720 | 3 | US-08-974-737-3 |
| 20 | 109 | 15.9 | 720 | 3 | US-08-706-408-3 |
| 21 | 109 | 15.9 | 720 | 3 | US-09-094-359-9 |
| 22 | 109 | 15.9 | 720 | 3 | US-09-172-063-12 |
| 23 | 109 | 15.9 | 720 | 3 | US-09-172-063-17 |
| 24 | 109 | 15.9 | 720 | 4 | US-09-465-142-3 |
| 25 | 109 | 15.9 | 720 | 4 | US-09-316-919-13 |
| 26 | 109 | 15.9 | 720 | 4 | US-09-316-919-18 |
| 27 | 109 | 15.9 | 720 | 4 | US-09-575-847-3 |

| | | | | | | |
|----|-----|------|------|---|-------------------|-------------------|
| 28 | 109 | 15.9 | 720 | 4 | US-09-602-641-12 | Sequence 12, Appl |
| 29 | 109 | 15.9 | 720 | 4 | US-09-602-641-17 | Sequence 17, Appl |
| 30 | 109 | 15.9 | 720 | 4 | US-09-704-463-9 | Sequence 9, Appl |
| 31 | 109 | 15.9 | 720 | 4 | US-09-316-920A-13 | Sequence 13, Appl |
| 32 | 109 | 15.9 | 720 | 4 | US-09-316-920A-18 | Sequence 18, Appl |
| 33 | 109 | 15.9 | 720 | 4 | US-10-071-976-3 | Sequence 3, Appl |
| 34 | 109 | 15.9 | 768 | 3 | US-09-172-063-28 | Sequence 28, Appl |
| 35 | 109 | 15.9 | 768 | 3 | US-09-172-063-33 | Sequence 33, Appl |
| 36 | 109 | 15.9 | 768 | 4 | US-09-602-641-28 | Sequence 28, Appl |
| 37 | 109 | 15.9 | 768 | 4 | US-09-602-641-33 | Sequence 33, Appl |
| 38 | 109 | 15.9 | 972 | 3 | US-09-172-063-30 | Sequence 30, Appl |
| 39 | 109 | 15.9 | 972 | 4 | US-09-602-641-30 | Sequence 30, Appl |
| 40 | 109 | 15.9 | 1623 | 4 | US-09-513-783A-33 | Sequence 33, Appl |
| 41 | 109 | 15.9 | 1623 | 4 | US-09-430-656-33 | Sequence 33, Appl |
| 42 | 109 | 15.9 | 2439 | 4 | US-09-513-783A-3 | Sequence 3, Appl |
| 43 | 109 | 15.9 | 2439 | 4 | US-09-513-783A-5 | Sequence 5, Appl |
| 44 | 109 | 15.9 | 2439 | 4 | US-09-430-656-3 | Sequence 3, Appl |
| 45 | 109 | 15.9 | 2439 | 4 | US-09-430-656-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-10-152-296-1
; Sequence 1, Application US/10152296
; Patent No. 6723537
; GENERAL INFORMATION:
; APPLICANT: Peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(723)
; OTHER INFORMATION: DsRED
US-10-152-296-1

Query Match 43.3%; Score 297.4; DB 4; Length 723;
Best Local Similarity 67.3%; Pred. No. 9.6e-53;
Matches 453; Conservative 0; Mismatches 211; Indels 9; Gaps 2;

| | | | |
|----|-----|---|-----|
| QY | 12 | CCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTA | 71 |
| Db | 21 | CGTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGCACCGTGAACGGCCACGA | 80 |
| QY | 72 | CTTCAAGTCCGAGGCGGAGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGAAAT | 131 |
| Db | 81 | GTTCGAGATCGAGGCGGAGGCGGAGGCGGCCCTTACGAGGGCCACAACACCCGTGAAGCT | 140 |
| QY | 132 | CCACGTACCGAGGCGGCCCTTCCTTCGCTTCGCTTCGACATCCTGGCCCCCTGTGGCA | 191 |
| Db | 141 | GAAGGTGACCAAGGCGGCCCTTCCTTCGCTTCGCTTCGACATCCTGGCCCCCTGTGGCA | 200 |
| QY | 192 | GTACGGCAGGAGGACCTTCGTGACCAACACCGCCGAGATCCCGGACTTCTTCAAGCAGAG | 251 |
| Db | 201 | GTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCGGACTACAAGAAGCTGTC | 260 |
| QY | 252 | CTTCCCGGAGGGCTTACCTGGGAGAGAACCAACCACTACGAGGACGGGGCATCTGTAC | 311 |

Db 261 CTTCCTCCGAGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAGGACGGCGCGTGGTGAC 320
QY 312 CGCCACACAGGACACAGCCCTGGAGGCAACTGCTGATCTACAAAGTGAAGGTGCACGG 371
Db 321 CGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCACTACAAAGTGAAGTTCATCGG 380
QY 372 CACCAACTTCCCGCGGACGGCCCGCTGATGAAGAAACAAGAGCGCGCTGGAGCCCGAG 431
Db 381 CGTGAACCTTCCCTCCGACGGCCCGCTAATGCAAGAAGAACCATGGGCTGGAGGCGCTC 440
QY 432 CACCGAGGTGGTGTACCCCGAGAACGGCGTGTGCTGCGGCGGAAACGTGATGGCCCTGAA 491
Db 441 CACCGAGCGCTGTACCCCGGACGGCGTGTGAGGGCGAGATCCACAAGGCCCTGAA 500
QY 492 GGTGGGCGACCGGCACCTGATCTGCCACCACTACACAGCTACCGGACGAAGAAGCGCT 551
Db 501 GCTGAAGGACGGCGCC-----ACTACCTGGTGGAGTTCAAGAGTATCTACATGGCCAA 554
QY 552 GCGCGCCTGACCATGCGCGCTTCCACTTCAACGACATCCGGCTCCAGATGCTGCGGAA 611
Db 555 GAAAGCCGTGACGTGCGCGCTACTACTACGTGGACTCCAAAGCTGGACATCACCTCCCA 614
QY 612 GAAGAAGGACGAGTACTTC---GAGCTGTACGAGGCGGAGCTGGCGGCTACAGCGACCT 668
Db 615 CAACGAGGACTACCATCGTGGAGCAGTACGAGCGCACCGAGGGCGCCACCACTGTT 674
QY 669 GCCCGAGAAGGCC 681
Db 675 CCGGAGGAGGCC 687

RESULT 2
US-09-977-897-1
; Sequence 1, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-1

Query Match 36.9%; Score 253.4; DB 4; Length 1482;
Best Local Similarity 63.1%; Pred. No. 1.2e-43;
Matches 427; Conservative 0; Mismatches 241; Indels 9; Gaps 2;
QY 2 TGGTGAAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCGACCGTGA 61
Db 32 TGAAGAACACCGGCTGAAGGAGATCATGAGCGGCAAGCGGCGTGGAGGCGATCGTGA 91
QY 62 ACGGCCACTACTTCAAGTGCAGGCGGAGGCGACCGCAACCCCTTCGCGGCGACCCAGA 121
Db 92 ACAACACAGTGTTCAGCATGAGGCGGCTTCGGCAAGGCAACAGTGTGTTCCGGCAACCCAGC 151
QY 122 GCATGAGAATCCACGTACCGAGGCGGCGCCCTTCGCTTCGCTTCGACATCCTGGCCC 181
Db 152 TGATGCAGATCCGGGTGACCAAGGCGGCGCTTCGCTTCGCTTCGACATCCTGGAGCA 211
QY 182 CCTGTGGAGTACGGCAGCAGGACCTTCGTGCAACCAACCGCGGAGATCCCGGACTTCT 241
Db 212 TCGCCTTCCAGTACGGCAACCGGACCTTCACCAAGTATCCCGACGACATCGCGGACTACT 271
QY 242 TCAAGCAGAGCTTCCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGAGCGG 301

Db 272 TCGTGACAGGCTTCCCTGCCGCTTCTTCTACAGCGGAACCTCGGTTTCGAGGACGGCG 331
QY 302 GCATCCTGACCGCCACAGGACACAGCCTGGAGGGCAACTGCTGATCTACAAGGTGA 361
Db 332 CCATCGTGGACATCGGAGCGACATAGCCTGGAGGACGACAAGTTCACACTACAAGGTGG 391
QY 362 AGGTGACGGCACCAACTTCCCGCGGACGGCCCGCTGATGAAGAAACAAGAGCGCGGCT 421
Db 392 AGTACCGCGGCAACGGCTTCCCTAGCAACGGCCCTGTGATGCAGAAGGCCATCCTGGGCA 451
QY 422 GGGAGCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGC---TGTGGGCGGGAACG 478
Db 452 TGGAGCCAGCTTCGAGGTGGTGTACATGAACAGCGCGTGTCTGGTGGCGAGGTGGACC 511
QY 479 TGATGGCCCTGAAGTGGGCGGACCGGCACTGTGATCTGCCACCACTACACAGCTACCGGA 538
Db 512 TGGTGTACAAGCTGGAGAGCGGCAACTACTACAGTGCACATGAAGACCTTCTACCGGA 571
QY 539 GCAAGAAGCGCGTGGCGCCCTGACCATGCGCGCTTCCACTTCAACGACATCCGGCTCC 598
Db 572 GCAAGGGCGGCGTG-----AAGGAGTTCCCTGAGTACCTATTCATCCACCAACCGGCTGG 625
QY 599 AGATGCTGCGGAAGAAGGACGAGTACTTCGAGTGTACGAGGCGGCGTGGCCCGGT 658
Db 626 AGAAGAACTACGTGGAGGAGGCGAGCTTCGTGGAGCAGCAGAGACCGCCATCGCCAGC 685
QY 659 ACAGCGACCTGCCCGAG 675
Db 686 TGACCACCATCGGCAAG 702

RESULT 3
US-09-839-650-1
; Sequence 1, Application US/09839650
; Patent No. 6645761
; GENERAL INFORMATION:
; APPLICANT: Stratagene
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green
; Patent No. 6645761
; TITLE OF INVENTION: Fluorescent Protein
; FILE REFERENCE: 25436/1755
; CURRENT APPLICATION NUMBER: US/09/839,650
; CURRENT FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanized R. mulleri polynucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)..(720)
; OTHER INFORMATION: Humanized DNA sequence
US-09-839-650-1

Query Match 36.9%; Score 253.2; DB 4; Length 720;
Best Local Similarity 62.0%; Pred. No. 1.2e-43;
Matches 418; Conservative 0; Mismatches 253; Indels 3; Gaps 1;
QY 2 TGGTGAAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCGACCGTGA 61
Db 20 TGAAGAACACCTGCTGCGAGGAGTGTAGCTACAAGGTGAACCTGGAGGCGATCGTGA 79
QY 62 ACGGCCACTACTTCAAGTGCAGGCGGCGAGGCGACCGCAACCCCTTCGCGGCGACCCAGA 121
Db 80 ACAACACAGTGTTCACCATGGAGGCGTGGCGCAAGGGCAACATCTGTTTCGGCAACCCAGC 139
QY 122 GCATGAGAATCCACGTACCGAGGCGGCGCCCTTCGCTTCGCTTCGACATCCTGGCCCC 181
Db 140 TGGTGCAGATCCGCGTGACCAAGGCGGCGCCCTTCGCTTCGCTTCGACATCGTGAGCC 199

QY 182 CCTGCTGCGAGTACGGCAGCAGGACCTTCTGTGACCAACACCGCCGAGATCCCCGACTTCT 241
|||
Db 200 CCGCTTCCAGTACGGCAACCGCACCTTCAACCAAGTACCCCAACGACATCAGCGACTACT 259
|||
QY 242 TCAAGCAGAGCTTCCCGAGGGCTTCACTCTGGGAGAGAACCAACCACTACGAGGACGGCG 301
|||
Db 260 TCATCCAGAGCTTCCCGCGGCTTCATGTACGAGCCACCCCTGCGCTACGAGGACGGCG 319
|||
QY 302 GATCTCTGACCGCCCAACAGGACACAGCCTTGAGGGCAACTGCTGTATCTACAAGGTGA 361
|||
Db 320 GCCTGGTGGAGATCCGAGCGACATCAACCTGTATCGAGGACAAGTTCTGTACTCCGCTGG 379
|||
QY 362 AGGTGCACGGCACCAACTTCCCGCCGACAGCGCCCGCTGATGAAGAACAAAGACGGCGCT 421
|||
Db 380 AGTACAAGGGCAGCAACTTCCCGACGACGCGCCCGTGTATGCAGAGAACCATCTCTGGGCA 439
|||
QY 422 GGGAGCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGA 481
|||
Db 440 TCGAGCCAGCTTCGAGGCCATGTACATGAACAACGGCGTGTGTGGCGAGGTGATCC 499
|||
QY 482 TGGCCCTGAAGTGGCGACCGGCACCTGATCTGCCACCACTACACCGACTACCGGAGCA 541
|||
Db 500 TGGTGTACAAGCTGAACAGCGCAAGTACTACAGCTGCCACATGAAGACCCCTGATGAAGA 559
|||
QY 542 AGAAGCGCGTGCAGCGCCCTGACCATGCCCGGCTTCCACTTCACCGACATCCGGCTCCAGA 601
|||
Db 560 GCAAGGGCGTG--GTGAAGGAGTTCCCTCCTACCACTTCATCCAGCACCGGCTGGAGA 616
|||
QY 602 TGCTGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACA 661
|||
Db 617 AGACTTACGTGGAGGACGGCGCTTCGTGGAGCAGCAGGACCGCCATCGCCACAGATGA 676
|||
QY 662 GCGACCTGCCCGAG 675
|||
Db 677 CCAGCATCGGCAAG 690
|||

RESULT 4

US-09-459-956-6
; Sequence 6, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp
US-09-459-956-6

Query Match 25.7%; Score 176.6; DB 3; Length 678;
Best Local Similarity 60.2%; Pred. No. 6.9e-28;
Matches 293; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 17 TGAAGGAGAGTATCGGCATCAAGATGTACATGGAGGGACCGTGAACGGCCACTTCA 76
|||
Db 23 TCAAGGAGTTTATGAGTTTAAAGTTTCGATGGAAGGAAACGGTCAATGGGACAGTTTG 82
|||
QY 77 AGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGGACCCAGAGCATGAGATCCACG 136
|||

Db 83 AATATAAGGCGAAGGAGAGGGAGGCCATACGAAGGCCACATACCGTAAAGCTTAAGG 142
|||
QY 137 TGACCGAGGGGCCCCCTTCCCTTTCGCCTTCGACATCTTGGCCCCCTGCTGCGAGTACG 196
|||
Db 143 TAACCAAGGGGGGACCTTTGCCATTTGCTTTGGGATATTTGTCAACCACAATTTCAATG 202
|||
QY 197 GCAGCAGGACCTTCTGTGCACCAACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCC 256
|||
Db 203 GAAGCAAGGTATATGTCAAGCACCTTCCGACATACCAACTATAAAAAGCTGTCTATTC 262
|||
QY 257 CCGAGGGCTTCACTTGGGAGAGAACCAACCACTTACGAGGACGGCGGCTCCTGACCGCCC 316
|||
Db 263 CTGAAGGATTTAAATGGGAAGGGTCAATGAACCTTTGAAGACGGTGGCGTCTGTTACTG 322
|||
QY 317 ACCAGGACACAGCCTTGAGGGCAACTGCCTGATCTACAAGGTGAAGGTGCACGGCACCA 376
|||
Db 323 CCCAGGATTCAGTTTGCAGGATGGCTGTTTTCATCTACAAGGTCAAGTTTCATTGGCGTGA 382
|||
QY 377 ACTTCCCGCCGACGGCCCCGTGATGAAGAACAAAGACGGCGGCTGGAGCCCAAGCACCG 436
|||
Db 383 ACTTTCCTTCCGATGGACCTTGTATGCAAAAGAAAGCAATGGGCTGGGAAGCCAGCACTG 442
|||
QY 437 AGGTGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGATGGCCCTGAAGGTGG 496
|||
Db 443 AGCGTTGTATCCTCGTGTGGCGTGTGTAAGAGGAGAGATTCTAAGGCTCTGAAGCTGA 502
|||
QY 497 GCGACCG 503
|||
Db 503 AAGACGG 509
|||

RESULT 5

US-09-459-956-7
; Sequence 7, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Clavularia sp
US-09-459-956-7

Query Match 22.4%; Score 153.8; DB 3; Length 801;
Best Local Similarity 58.4%; Pred. No. 3.5e-23;
Matches 269; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 10 GGCCTGCTGAAGGAGAGTATGCCCATCAAGATGTACATGGAGGGCACCGTGAACGGCCAC 69
|||
Db 130 GGTGTGATTAAACCAACGACATGAAGATTAAAGCTGAAGATGAAGGAATGTAAACGGGCAT 189
|||
QY 70 TACTTCAAGTCCGAGGGCGAGGGGACGGCAACCCCTTCGCGCGCACCCAGAGCATGAGA 129
|||
Db 190 GCTTTTGTGATCGAAGGAGAGAGAGAAAGCCCTTACGATGGGACACACACTTTAAAC 249
|||
QY 130 ATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGCTTCGACATCTTGGCCCCCTGCTGC 189
|||
Db 250 CTGGAAGTCAAGGAAGGTGCGCTCTGCCTTTTCTTACGATATCTTGTCAAACCGGCTC 309
|||

QY 376 AACTTCCCGCGACGGCCCCGTGATGAAGAAACAAGAGCGGCGGTGGAGCCCGACACC 435
|||||
Db 424 GGTTCCTAGTAACGACCCCGTGTGATGCAAAAAGCCATCCTCGGCATGGAGCCATCGTTT 483
|||||
QY 436 GAGGTGGTGTACCCCGAGAACGGCGGTGCTGTGCGGC 471
|||||
Db 484 GAGGTGGTGTACATGAACAGCGGCGTCTGTGTGGGC 519
|||||

RESULT 8

US-09-609-161B-30
; Sequence 30, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(747)
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert A)
US-09-609-161B-30

Query Match 17.7%; Score 121.6; DB 4; Length 1104;
Best Local Similarity 54.2%; Pred. No. 1.5e-16;
Matches 247; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
QY 16 CTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGACCGTGAACGGCCACTACTTC 75
|||||
Db 64 CTGAAGAGATTATGTGCGCAAAAGCTAGCGTTGAAGGAATCGTGAACAATCACGTTTTT 123
|||||
QY 76 AAGTGGAGGGCGAGGGCGACGCAACCCCTTCGCCGGCACCAGAGCATGAGATCCAC 135
|||||
Db 124 TCCATGGAAGGATTGTGGAAGGCAATGTATTATTGGAACCAATTGATGCAATCCGG 183
|||||
QY 136 GTGACCGAGGGCGCCCCCTTCGCCCTTCGACATCCTGGCCCCCTGCTGGGAGTAC 195
|||||
Db 184 GTTACAAAGGGAGGTCGGTTGCCATTGCTTTCGATATTGTTCCATAGCTTTCATATAC 243
|||||
QY 196 GGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTCTTCAAGCAGAGCTTC 255
|||||
Db 244 GGGATCGCACTTTTCAGAAATACCCAGACGACATTGCGGACTACTTTGTTCAATCATTC 303
|||||
QY 256 CCCGAGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGCGGCATCCTGACCGCC 315
|||||
Db 304 CCGGCTGGATTTTTCTACGAAAGAAATCTACGCTTTGAAGATGGCGCCATTGTTGACATT 363
|||||
QY 316 CACCAAGACACCGCTGGAGGCAACTGCTGATCTACAAGGTGAAGGTGACGGCACC 375
|||||
Db 364 CGTTCAGATATAAGTTTAGAAGATGATAAGTTCCTACTACAAAGTGGAGTATAGAGGCAAC 423
|||||
QY 376 AACTTCCCGCGACGGCCCCGTGATGAAGAAACAAGAGCGGCGGTGGAGCCCGACACC 435
|||||

Db 424 GGTTCCTAGTAACGACCCCGTGTGATGCAAAAAGCCATCCTCGGCATGGAGCCATCGTTT 483
|||||
QY 436 GAGGTGGTGTACCCCGAGAACGGCGGTGCTGTGCGGC 471
|||||
Db 484 GAGGTGGTGTACATGAACAGCGGCGTCTGTGTGGGC 519
|||||

RESULT 9

US-09-277-716-31
; Sequence 31, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(720)
; FEATURE:
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert B)
US-09-277-716-31

Query Match 17.7%; Score 121.6; DB 3; Length 1279;
Best Local Similarity 54.2%; Pred. No. 1.6e-16;
Matches 247; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
QY 16 CTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGACCGTGAACGGCCACTACTTC 75
|||||
Db 37 CTGAAGAGATTATGTGCGCAAAAGCTAGCGTTGAAGGAATCGTGAACAATCACGTTTTT 96
|||||
QY 76 AAGTGGAGGGCGAGGGCGACGCAACCCCTTCGCCGGCACCAGAGCATGAGATCCAC 135
|||||
Db 97 TCCATGGAAGGATTGTGGAAGGCAATGTATTATTGGAACCAATTGATGCAATCCGG 156
|||||
QY 136 GTGACCGAGGGCGCCCCCTTCGCCCTTCGACATCCTGGCCCCCTGCTGGGAGTAC 195
|||||
Db 157 GTTACAAAGGGAGGTCGGTTGCCATTGCTTTCGACATTGTTCCATAGCTTTCATATAC 216
|||||
QY 196 GGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTCTTCAAGCAGAGCTTC 255
|||||
Db 217 GGGATCGCACTTTTCAGAAATACCCAGACGACATTGCGGACTACTTTGTTCAATCATTT 276
|||||
QY 256 CCCGAGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGCGGCATCCTGACCGCC 315
|||||
Db 277 CCGGCTGGATTTTTCTACGAAAGAAATCTACGCTTTGAAGATGGCGCCATTGTTGACATT 336
|||||
QY 316 CACCAAGACACCGCTGGAGGCAACTGCTGATCTACAAGGTGAAGGTGACGGCACC 375
|||||
Db 337 CGTTCAGATATAAGTTTAGAAGATGATAAGTTCCTACTACAAAGTGGAGTATAGAGGCAAC 396
|||||
QY 376 AACTTCCCGCGACGGCCCCGTGATGAAGAAACAAGAGCGGCGGTGGAGCCCGACACC 435
|||||
Db 397 GGTTCCTCCCTAGTAACGGACCCCGTATGCAAAAAGCCATCCTCGGCATGGAGCCATCGTTT 456
|||||
QY 436 GAGGTGGTGTACCCCGAGAACGGCGGTGCTGTGCGGC 471
|||||
Db 457 GAGGTGGTGTACATGAACAGCGGCGTCTGTGTGGGC 492
|||||

RESULT 10
US-09-609-161B-31
; Sequence 31, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(720)
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert B)
US-09-609-161B-31

Query Match 17.7%; Score 121.6; DB 4; Length 1279;
Best Local Similarity 54.2%; Pred. No. 1.6e-16;
Matches 247; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
QY 16 CTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGCCACTACTTC 75
Db 37 CTGAAGGAGATTATGTCGCAAAAGCTAGCGTTGAAGGAATCGTGAACAATCACGTTTTT 96
QY 76 AAGTCGAGGGCGAGGGCGACGGCAACCCCTTCGCGGCACCCAGAGCATGAGAATCCAC 135
Db 97 TCCATGGAAGGATTGGAAGGCAATGTAATTTGGAACCAATGATGCAATCCGG 156
QY 136 GTGACCGAGGGCGCCCTCGCCCTTCGCTTCGACATCCTGGCCCTCGCTCGAGTAC 195
Db 157 GTTACAAAGGGAGGTCCGTTGCCATTGCTTTTCGACATGTTTCCATAGCTTCCAATAC 216
QY 196 GGCAGCAGGACCTTCGTGCACCAACACCGCCGAGATCCCGACTTCTTCAAGCAGAGCTTC 255
Db 217 GCGAATCGCACTTTCACGAAATACCCAGACACATTCGCGACTACTTTGTTCAATCATTT 276
QY 256 CCGAGGGCTTCACCTGGAGAGAACCAACCACTACGAGGACGGCGGCATCCTGACCGCC 315
Db 277 CCGCTGGATTTTCTACGAAAGAAATCTACGCTTGAAGATGGCGCCATTGTTGACATT 336
QY 316 CACCAGGACACCGCCCTGGAGGGCAACTGCTGTATCTACAAGGTGAAGTGCACGGCACC 375
Db 337 CGTTCAGATATAAGTTTGAAGATGATAAGTTTCCACTACAAAGTGGAGTATAGAGGCAAC 396
QY 376 AACTTCCCGCCGACGGCCCGCTGATGAAGAACAGAGCGCGCTGGAGCCCGCAGCACC 435
Db 397 GGTTCCTTAGTAACGGACCGCTGATGTCAAAAGCCCATCCTCGGCATGGAGCCATCGTTT 456
QY 436 GAGGTGGTGTACCCCGAGAACGGCGTGTCTGTGGGC 471
Db 457 GAGGTGGTCTACATGAACAGCGGCGTCTGTGTGGGC 492

RESULT 11

US-09-172-063-14
; Sequence 14, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: EYFP-V68L/Q69K
US-09-172-063-14
Query Match 16.1%; Score 110.6; DB 3; Length 720;
Best Local Similarity 55.2%; Pred. No. 2.7e-14;
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;
QY 13 CTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGCCACTAC 72
Db 22 CTGTTACCGGGGTGTGTCCTCATCTGTCGAGTGGACGGCGACGTAACGGCCACAAG 81
QY 73 TTCAAGTGGAGGGCGAGGGCGACGGCAACCCCTTCGCGGCACCCAGAGCATGAGAATC 132
Db 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGTCACCTACGGCAAGCTGACCTGAAGTTC 141
QY 133 CACGTGACCGAGGGCGCCCTTCGCTTCGCTTCGACATCCTGGCCCTGTCGCGAG 192
Db 142 ATCTGCACCAC---CGCAAGCTGCCCTGCTGCCCCACCCCTGTCGACCACTTCGGC 198
QY 193 TACGGCAGCAGGACCTTCGTGCACCAACACCGCCGAGAT-----CCCCGACTTCTTCAAG 246
Db 199 TACGGCCTGAAGTGTTCGCCCCGTACCCCGACCAACATGAAGCAGCAGCACTTCTTCAAG 258
QY 247 CAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGACGGCGCATC 306
Db 259 TCCGCCATGCCGAGAGGTACGTCCAGGAGGCGACCATCTTCTTCAAGGACGCGCAAC 318
QY 307 CTGACCGCCCAACAGGACACCAAGCTGGAGGGCAACTGCTGATCTACAAGGTGAAGTG 366
Db 319 TACAAGACCCCGCGAGGTGAAGTTCGAGGGCGACACCCCTGTTGACCGCATCGAGCTG 378
QY 367 CACGGCACCAACTTCCCGCCGACCGGCCCTGTGATGAAGAACAGAGCGCGCTGGGAG 426
Db 379 AAGGCGATCGACTTCAAGGAGGACGGCAACATCTTGGGGCAACAAGCTGGAGTACAATAC 438
QY 427 CCCAGCACCGAGGTGTGTATACCCGAGAACGGCGTGTGTGCGGCCGGAACGTGA 481
Db 439 AACAGCCCAACCGTCTATATCATGGCCGACAGCAGAGAACGGCATCAAGGTGA 493

RESULT 12
US-09-316-919-15
; Sequence 15, Application US/09316919
; Patent No. 6469154
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS


```
; FILE REFERENCE: 07257/073001
; CURRENT APPLICATION NUMBER: US/09/316,919
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
US-09-316-919-15

Query Match      16.1%; Score 110.6; DB 4; Length 720;
Best Local Similarity 55.2%; Pred. No. 2.7e-14;
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;

QY 13 CTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTAC 72
Db 22 CTGTTCAACCGGGTGGTGCCCATCTCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG 81
QY 73 TTCAAGTGGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGAATC 132
Db 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTC 141
QY 133 CACGTGACCGAGGGCGCCCGCTGCCCTTCGCCTTCGACATCTCGGCCCGCTGTCGGAG 192
Db 142 ATCTGCACCAC--CGGCAAGCTGCCCTGCCCTGCCCGACCATCTCTTCAAG 198
QY 193 TACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGAT-----CCCCGACTTCTTCAAG 246
Db 199 TACGGCCTGAAGTGCTTCGCCCGGTACCCCGACCATGAGCAGCAGCACTTCTTCAAG 258
QY 247 CAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCGCATC 306
Db 259 TCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAACATCTTCTCAAGGACGACGCAAC 318
QY 307 CTGACCGCCACACGAGACACACGCTGGAGGGCAACTGCCTGATCTACAAGGTGAGGTG 366
Db 319 TACAAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTG 378
QY 367 CACGGCACCAACTTCCCGCGGACGGCCCGCTGATGAAGAACAGAGCGGCGTGGAG 426
Db 379 AAGGGCATCGACTTCAAGGAGGACGGCAACATCTTGGGGCAAGAGCTGGAGTACAAC 438
QY 427 CCCAGCACCGAGGTGGTGTATCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGA 481
Db 439 AACAGCCACAACGTCTATATCATGCGCGACAGCAAGAACGCGCATCAAGGTGA 493
```

```
RESULT 13
US-09-602-641-14
; Sequence 14, Application US/09602641
; Patent No. 6608189
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wächter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/602,641
; CURRENT FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (0)...(0)
; OTHER INFORMATION: EYFP-V68L/Q69K
US-09-602-641-14

Query Match      16.1%; Score 110.6; DB 4; Length 720;
Best Local Similarity 55.2%; Pred. No. 2.7e-14;
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;

QY 13 CTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTAC 72
Db 22 CTGTTCAACCGGGTGGTGCCCATCTCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG 81
QY 73 TTCAAGTGGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGAATC 132
Db 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTC 141
QY 133 CACGTGACCGAGGGCGCCCGCTGCCCTTCGCCTTCGACATCTCGGCCCGCTGTCGGAG 192
Db 142 ATCTGCACCAC--CGGCAAGCTGCCCTGCCCTGCCCGACCATCTCTTCAAG 198
QY 193 TACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGAT-----CCCCGACTTCTTCAAG 246
Db 199 TACGGCCTGAAGTGCTTCGCCCGGTACCCCGACCATGAGCAGCAGCACTTCTTCAAG 258
QY 247 CAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCGCATC 306
Db 259 TCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAACATCTTCTCAAGGACGACGCAAC 318
QY 307 CTGACCGCCACACGAGACACACGCTGGAGGGCAACTGCCTGATCTACAAGGTGAGGTG 366
Db 319 TACAAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTG 378
QY 367 CACGGCACCAACTTCCCGCGGACGGCCCGCTGATGAAGAACAGAGCGGCGTGGAG 426
Db 379 AAGGGCATCGACTTCAAGGAGGACGGCAACATCTTGGGGCAAGAGCTGGAGTACAAC 438
QY 427 CCCAGCACCGAGGTGGTGTATCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGA 481
Db 439 AACAGCCACAACGTCTATATCATGCGCGACAGCAAGAACGCGCATCAAGGTGA 493
```

```
RESULT 14
US-09-316-920A-15
; Sequence 15, Application US/09316920A
; Patent No. 6699687
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470
; CURRENT APPLICATION NUMBER: US/09/316,920A
; CURRENT FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
US-09-316-920A-15

Query Match      16.1%; Score 110.6; DB 4; Length 720;
Best Local Similarity 55.2%; Pred. No. 2.7e-14;
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;

QY 13 CTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTAC 72
Db 22 CTGTTCAACCGGGTGGTGCCCATCTCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG 81
QY 73 TTCAAGTGGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGAATC 132
Db 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCCTGAAGTTC 141
```

Qy 133 CACGTGACCGAGGGCGCCCTTCCCTTCGCTTCGACATCCTGGCCCCCTGCTGGAG 192
Db 142 ATCTGCACCAAC---CGGCAAGCTGCCCTGCTGCCCTTGGCCCCACCTCGTGACCACTTCGGC 198
Qy 193 TACGGCAGCAGGACCTTCGTGCACCAACACACCGCCGAGAT-----CCCCGACTTCTTCAAG 246
Db 199 TACGGCCTGAAGTGCTTCGCCCGCTACCCCGACCAATGAGCAGCAGCACTTCTTCAAG 258
Qy 247 CAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGACGGCGGCATC 305
Db 259 TCCGCCATGCCCGAAGGTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAAC 318
Qy 307 CTGACCGCCCAACAGGACACCAAGCTGGAGGGCACTGCCTGATCTACAAGGTGAAGTG 365
Db 319 TACAAGACCCGCGCGAGGTGAAGTTCGAGGGCGCACACCTTGGTGAACCCGCATCGAGCTG 378
Qy 367 CACGGCACCAACTTCCCGCGCAGCGGCCCTGTGATGAAGAACAAAGAGCGGGCGCTGGAG 426
Db 379 AAGGCATCGACTTCAAGGAGGACGGCAACATCTGGGGCAACAAGCTGGAGTACAACCTAC 438
Qy 427 CCAGCACCGAGGTGTGTACCCCGAGAACCGGCTGCTGTGCGGCCGGAACGTGA 481
Db 439 AACAGCCACAACGTCTATATCATGTGGCCGCAAGCAGAAAGAACGGCATCAAGGTGA 493

RESULT 15

US-09-459-956-2
; Sequence 2, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Anemonia majano
US-09-459-956-2

Query Match 16.0%; Score 109.8; DB 3; Length 690;
Best Local Similarity 57.0%; Pred. No. 3.9e-14;
Matches 223; Conservative 0; Mismatches 162; Indels 6; Gaps 1;
Qy 44 ACATGGAGGGCACCGTGAACGGCCACTTCAAGTGCAGGGCGAGGGCGACGGCAACC 103
Db 50 ATATGGATGGCTGTGTCAATGGGCATTACTTTACCGTCAAAGGTGAAGGCAACGGGAAGC 109
Qy 104 CCTTCGCGCGCACCCAGAGCATGAG-----AATCCACGTGACCGAGGGCGCCCCCTGC 157
Db 110 CATACGAAGGGACGCGAGACTTCGACTTTTAAAGTCACCATGGCCAACGGTGGGCCCTTG 169
Qy 158 CCTTCGCGCTTCGACATCCTGCGCCCCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACC 217
Db 170 CATTCCTCTTGACATACATCTACAGTGTTCATAATATGGAATCGATGCTTTACTGCGT 229
Qy 218 ACACCGCGGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGA 277
Db 230 ATCTTACCAGTATGCCCGACTATTTCAAACAAGCATTTCTCTGACGGAATGTCATATGAA 289
Qy 278 GAACCAACCACTACGAGGAGCGGGCGCATCCTGACCGCCCCACCAGGACACCAAGCCTGGAGG 337

Db 290 GGACTTTTACCTATGAAGATGGAGGAGTTGCTACAGCCAGTTGGGAAATAAGCCTTAAAG 349
Qy 338 GCAACTGCCTGATCTACAAGGTGAAGGTGCACGCGCACCAACTTCCCCCGCGACGGCCCCG 397
Db 350 GCAACTGCTTTGAGCCACAATAATCCACGTTTTCATGGAGTGAACCTTTCCTGCTGATGGACCTG 409
Qy 398 TGATGAAGAACAAGAGCGCGGCTGGGAGCC 428
Db 410 TGATGGCGAAGAAGACAACACTGGTTGGGACCC 440

Search completed: December 10, 2004, 17:42:19
Job time : 102 secs

; SEQ ID NO 15
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion construct
US-09-976-673-15

Query Match 97.7%; Score 671; DB 9; Length 1396;
Best Local Similarity 98.5%; Pred. No. 9.3e-163;
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCAACCGTG 60
Db |||||
QY 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120
Db |||||
QY 121 AGCATGAGATCCACGTGACCGAGGGCGCCCCCTGCCCTTCGACATCCTGGCC 180
Db |||||
QY 181 CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCCGACTTC 240
Db |||||
QY 241 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTACGAGGACGGC 300
Db |||||
QY 301 GGCATCTGTACCGCCCCACAGGACACAGCCTGGAGGGAACCTGCCTGATCTACAAGGTG 360
Db |||||
QY 361 AAGGTGACGGCACCAACTTCCCCGCGGACGGCCCTGATCTGCCACCACTACACGAGTACCGGAGC 420
Db |||||
QY 421 TGGAGCCCGAGCAGGAGGTGTGTACCCCGAGAACCGCGTGTGTGCGGCCGGAACGTG 480
Db |||||
QY 481 ATGGCCCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACGAGTACCGGAGC 540
Db |||||
QY 541 AAGAGCCGTGCGCGCCCTGACCATGCCGGCTTCCACTTCAACGACATCCGGTCCAG 600
Db |||||
QY 601 ATGCTGGGAAGGAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db |||||
QY 661 AGCGACCTGCCCGAGAGGCCCACTGA 687
Db |||||
QY 1370 AGCGACCTGCCCGAGAGGCCCACTGA 1396

RESULT 4
US-10-806-930-1
; Sequence 1, Application US/10806930
; Publication No. US20040216180A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; TITLE OF INVENTION: Nucleic Acids Encoding Linked
; FILE OF INVENTION: Chromo/Fluorescent Domains and Methods for Using the Same
; FILE REFERENCE: CLON-094
; CURRENT APPLICATION NUMBER: US/10/806,930
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: 09/976,673

; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/356,225
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/383,336
; PRIOR FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: PCT/US02/32560
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Anthozoa
US-10-806-930-1

Query Match 97.7%; Score 671; DB 18; Length 1396;
Best Local Similarity 98.5%; Pred. No. 9.3e-163;
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCAACCGTG 60
Db |||||
QY 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120
Db |||||
QY 121 AGCATGAGATCCACGTGACCGAGGGCGCCCCCTGCCCTTCGACATCCTGGCC 180
Db |||||
QY 181 CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCCGACTTC 240
Db |||||
QY 241 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTACGAGGACGGC 300
Db |||||
QY 301 GGCATCTGTACCGCCCCACAGGACACAGCCTGGAGGGAACCTGCCTGATCTACAAGGTG 360
Db |||||
QY 361 AAGGTGACCGCACCAACTTCCCCGCGGACGGCCCTGATGAAGAACAGAGCGGCGGC 420
Db |||||
QY 421 TGGAGCCCGAGCAGGAGGTGTGTACCCCGAGAACCGCGTGTGTGCGGCCGGAACGTG 480
Db |||||
QY 481 ATGGCCCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACGAGTACCGGAGC 540
Db |||||
QY 541 AAGAGCCGTGCGCGCCCTGACCATGCCGGCTTCCACTTCAACGACATCCGGTCCAG 600
Db |||||
QY 601 ATGCTGGGAAGGAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db |||||
QY 661 AGCGACCTGCCCGAGAGGCCCACTGA 687
Db |||||
QY 1370 AGCGACCTGCCCGAGAGGCCCACTGA 1396

RESULT 5
US-09-976-673-17
; Sequence 17, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:

Db 554 AAGAGGCGGTGCGGGCCCTGACCATGCCCGGCTTCCACTTACCGACATCCGGCTGCAG 613
Qy 601 ATGCTGCGGAAGAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db 614 ATGCTGCGGAAGGAGAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 673
Qy 661 AGCGACTGCCCGGAGAGGCCAACTGA 687
Db 674 AGCGACTGCCCGGAGAGGCCAAACAGA 700

RESULT 7
US-10-439-262-12
; Sequence 12, Application US/10439262
; Publication No. US20040106566A1
; GENERAL INFORMATION:
; APPLICANT: LIN, Shi-Lung
; YING, Shao-Yao
; TITLE OF INVENTION: RNA-Splicing and Processing-Directed
; Gene Silencing and The Relative Applications Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIN, Shi-Lung
; STREET: 1953 Wellesley Road,
; CITY: San Marino
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 91108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/439,262
; FILING DATE: 15-May-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/381,651
; FILING DATE: 17-MAY-2002
; APPLICATION NUMBER: US 60/411,062
; FILING DATE: 16-SEP-2002
; APPLICATION NUMBER: US 60/418,405
; FILING DATE: 12-OCT-2002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626)289-7172
; TELEFAX: (323)442-3158
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: mutated red fluorescent protein (rGFP) gene
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-439-262-12

Query Match 84.9%; Score 583; DB 17; Length 649;
Best Local Similarity 93.8%; Pred. No. 3.6e-140;
Matches 646; Conservative 0; Mismatches 0; Indels 43; Gaps 2;

Qy 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCGACCGGTG 60
Db 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGCGACCGGTG 60
Qy 61 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120
Db 61 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120

Qy 121 AGCATGAGAATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGACATCCTGGCC 180
Db 121 AGCATGAGAATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGACATCCTGGCC 180
Qy 181 CCCTGCTGCGAGTACGGCAGCAG---GACCTTCGTGACACACCGCCGAGATCCCGAC 237
Db 181 CCCTGCTGCGAGTACGGCAGCAGCAGCCTTCGTGACACACCGCCGAGATCCCGAC 240
Qy 238 TTCTTCAAGCAGAGCTTCCCGAGGGCTTACCTGGGAGAGAACCAACCACTACGAGGAC 297
Db 241 TTCTTCAAGCAGAGCTTCCCGAGGGCTTCACTGGGAG----- 279
Qy 298 GGCGGCATCCTGACCGCCCAACAGGACACACAGCCTGGAGGGCAACTGCCTGATCTACAAG 357
Db 280 -----ACCAAGGACACAGCCTGGAGGGCAACTGCCTGATCTACAAG 320
Qy 358 GTGAAGGTGACCGCACCAACTTCCCGCCGACCGGCCCGCTGATGAAGAAACAAGAGCGGC 417
Db 321 GTGAAGGTGACCGCACCAACTTCCCGCCGACCGGCCCGCTGATGAAGAAACAAGAGCGGC 380
Qy 418 GGCTGGGAGCCCGACGACCGAGGTGGTGTACCCCGAGAACGGCGTGTGCGGCCGGAAC 477
Db 381 GGCTGGGAGCCCGACGACCGAGGTGGTGTACCCCGAGAACGGCGTGTGCGGCCGGAAC 440
Qy 478 GTGATGSCCCTGAAGGTGGCGACCGGCACCTGATCTGCCACCACTACACAGTACCGG 537
Db 441 GTGATGSCCCTGAAGGTGGCGACCGGCACCTGATCTGCCACCACTACACAGTACCGG 500
Qy 538 AGCAAGAAGSCCGTGGCGCCCTGACCATGCCCGCTTCCACTTCCCGGACATCCGGCTC 597
Db 501 AGCAAGAAGSCCGTGGCGCCCTGACCATGCCCGCTTCCACTTCCCGGACATCCGGCTC 560
Qy 598 CAGATGCTCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACAGGCCAGCGTGGCCCCGG 657
Db 561 CAGATGCTCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACAGGCCAGCGTGGCCCCGG 620
Qy 658 TACAGCGACCTGCCCGAGAGGCCCAACTG 686
Db 621 TACAGCGACCTGCCCGAGAGGCCCAACTG 649

RESULT 8
US-10-806-930-5
; Sequence 5, Application US/10806930
; Publication No. US20040216180A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; TITLE OF INVENTION: Nucleic Acids Encoding Linked
; TITLE OF INVENTION: Chromo/Fluorescent Domains and Methods for Using the Same
; FILE REFERENCE: CLON-094
; CURRENT APPLICATION NUMBER: US/10/806,930
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: 09/976,673
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/356,225
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/383,336
; PRIOR FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: PCT/US02/32560
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Anthozoa
US-10-806-930-5

Query Match 55.0%; Score 446.6; DB 18; Length 1376;
Best Local Similarity 78.2%; Pred. No. 4.2e-105;
Matches 536; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
Qy 3 GGTGAGCGGCTGTGAAGGAGAGTATGCCATCAAGATGTACATGGAGGGCACCGTGAA 62

Db 692 GATGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGCAGCGTTAA 751
QY 63 CGGCCACTACTTCAAGTGCAGGGGAGGGCGACGGCAACCCCTTCGCGGACCCAGAG 122
Db 752 TGGCCATTATTTCAAGTGTGAAGAGAGGGAGACGGCAACCCATTTCAGAGTACGAG 811
QY 123 CATGAGATCCACGTGACCGAGGGGCGCCCTTCGCTTCGCTTCGACATCCCTGGCCCC 182
Db 812 CATGAGATTCATGTCAACCGAGGGGCTCCATTACCATTTGCTTCGACATTTGGCACC 871
QY 183 CTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCAACCGCGGAGATCCCGACTTCTT 242
Db 872 GTGTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCCGATTCTT 931
QY 243 CAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCG 302
Db 932 CAAGCAGCTTTCCTGAAGGCTTACTTTGGGAAAGAACCAACCTATGAAGATGGAG 991
QY 303 CATCTGACCGCCCAACAGGACACAGCCTGGAGGGCAACTGCCTGATCTCAAGGTGAA 362
Db 992 CATCTTACTGCTCATCAGGACACAGCCTGGAGGGCAACTGCCTTATATACAGGTGAA 1051
QY 363 GGTGACCGCACCAACTTCCCGCGGACCGGCGGCTGATGAAGAACAGAGCGGCGGTG 422
Db 1052 AGTCCATGGTACCAATTTTCTGCTGATGGCCCGGCTGATGAAGAACAAATCAGGAGGATG 1111
QY 423 GGAGCCAGCACCGAGGTGTTACCCCGAGAACGGCGTGTGCGGCGGCAACGTGAT 482
Db 1112 GGAGCCAGCACTGAGGTGTTATCCAGAGAAATGGTGTCTGTGAGGACGTAATGTGAT 1171
QY 483 GGCCTGAAGGTGGCGACCGGCACTGATCTGCACCACTACACAGCTACCGGAGCAA 542
Db 1172 GGCCTTAAAGTCGGTATCGTCATTTGATCTGCCATCACTATATCTTCTACAGGTCCAA 1231
QY 543 GAAGCGCGTGGCGCCCTGACCATGCCGCTTCCAGTTCACCGACATCCGCGCTCCAGAT 602
Db 1232 GAAGCAGTCCGTGCTTGCATATGCCAGGATTTCAATTTACAGACATCCGCTTCAGAT 1291
QY 603 GTCGCGGAAGAAAGGACAGTACTTCGAGCTGACGAGGCGCAGCGTGGCCCGGTACAG 662
Db 1292 GCTGAGGAAAAAGAAAGACAGTACTTTGAACCTGACGAAGCATCTGTGGTAGGTACAG 1351
QY 663 CGACCTGCCGAGAAAGCCAACTGA 687
Db 1352 TGATCTTCTGAAAAAGCAAAATTGA 1376

RESULT 9

US-09-976-673-7
; Sequence 7, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Fradkov, Arcady
; APPLICANT: Labas, Yulii
; APPLICANT: Matz, Mikhail
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Gurskaya, Nadezda
; TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
; FILE REFERENCE: CLON-028WO
; CURRENT APPLICATION NUMBER: US/09/976,673
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,018
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60,306,131
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 681
; TYPE: DNA
; ORGANISM: heteractis crispa

US-09-976-673-7
Query Match 64.2%; Score 441.2; DB 9; Length 681;
Best Local Similarity 78.2%; Pred. No. 9.8e-104;
Matches 530; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
QY 10 GGCCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTGAACGGCCAC 69
Db 4 GGTGTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATGGCCAT 63
QY 70 TACTTCAAGTGCAGGGGCGAGGGCGACGGCAACCCCTTCGCGGCAACCCAGAGCATGAGA 129
Db 64 TATTTCAAGTGTGAAGGAGAGGGGAGACGGCAACCCATTTCAGGTACGAGAGCATGAGG 123
QY 130 ATCCACGTACCGAGGGGCGCCCTTCGCTTCGCTTCGACATCCTGGCCCTGCTGC 189
Db 124 ATTCATGTACCGAAGGGGCTCCATTACCATTTGCTTCGACATTTTGGCAGCGTGTGT 183
QY 190 GAGTACGGCAGCAGGACCTTCGTGCACCAACACCGCGGAGATCCCGACTTCTTCAAGCAG 249
Db 184 GAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCCGATTCTTCAAGCAG 243
QY 250 AGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACACCTACGAGGACGGCGGCGCATCTG 309
Db 244 TCTTTCCTGAAGGCTTACTTTGGGAAAGAACCAACCACTATGAAGATGAGGCGCATCTT 303
QY 310 ACCGCCACCACTTCCCGCGACCGCTGGAGGGCAACTGCCTGATCTACAAGTGAAGGTGCAC 369
Db 304 ACTGCTCATCAGGACACAAAGCCTGGAGGGAACTGCTTATATACAAGTGAAGTCCCTT 363
QY 370 GGCACCAACTTCCCGCGACCGGCGCCCTGATGAAGAACAAAGAGCGGCGGCTGGAGCCCC 429
Db 364 GGTACCAATTTTCTGCTGATGGCCCGGCTGATGAAGAACAAATCAGGAGGATGGAGCCCA 423
QY 430 AGCACCGAGGTGTTTATCCAGAGAAATGGTGTCTGTGCGGCGGAACTGATGGCCCTG 489
Db 424 AGCACTGAGGTGTTTATCCAGAGAAATGGTGTCTGTGCGGCGGAACTGATGGCCCTT 483
QY 490 AAGTGGCGGACCGGCACTGATCTGCCACCACTACACAGCTACCGGAGCAAGAGGCC 549
Db 484 AAAGTCGGTGATCGTCTGTTGATCTGCCATCACTATACCTTCTTACAGTCCCAAGAAAGCA 543
QY 550 GTCGCGCCCTGACCATGCCGCTTCCACTTCCAGCATCCGCTCCAGATGCTGCGG 609
Db 544 GTCCGTGCTTGACAAATGCCAGGATTTCAATTTACAGACATCCGCTTCAGATGCTGAGG 603
QY 610 AAGAAAGAGGACGAGTACTTCGAGCTGTACGAGGCGGCGGCTGGCCCGGTACAGCGACCTG 669
Db 604 AAAGAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTGTCTT 663
QY 670 CCCGAGAAAGGCCAACTGA 687
Db 664 CCTGAAAAAGCAAAATTGA 681

RESULT 10

US-09-976-673-9
; Sequence 9, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Fradkov, Arcady
; APPLICANT: Labas, Yulii
; APPLICANT: Matz, Mikhail
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Gurskaya, Nadezda
; TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
; FILE REFERENCE: CLON-028WO
; CURRENT APPLICATION NUMBER: US/09/976,673
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,018
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60,306,131


```

; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 681
; TYPE: DNA
; ORGANISM: heteractis crispa
US-09-976-673-9

Query Match      64.0%; Score 439.6; DB 9; Length 681;
Best Local Similarity 78.0%; Pred. No. 2.5e-103;
Matches 529; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 10 GGCCTGCTCAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAAACGGCCAC 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 GGTGTTGTAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATGGCCAT 63

QY 70 TACTTCAAGTCGAGGGCGAGGGCGACGGCAACCCCTTCGCGCGCACCCAGAGCATGAGA 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 TATTTCAAGTGTGAAGAGAGAGGGAGACGGCAACCCATTTGCAGGTACGCAGAGCATGAGG 123

QY 130 ATCCACGTGACCGAGGGCGCCCTTCGCTGCCCTTCGCTTCGACATCCTGGCCCCCTGCTGC 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 ATTCATGTACCCGAAGGGGCTCCATTACCATTTGCCCTTCGACATTTTGGCACCCGTGTTGT 183

QY 190 GAGTACGGCAGCAGGACCTTCGTGCACCAACACCGCCGAGATCCCGACTTCTTCAAGCAG 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 GCGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCTCGATTTCTTCAAGCAG 243

QY 250 AGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCGGCACTCCTG 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 TCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGGAGGCACTTCT 303

QY 310 ACCGCCACACAGGACACAGCCTGGAGGGCAACTGCCCTGATCTACAAGGTGAAGTGCAC 369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 ACTGCTCATCAGGACACAAAGCCTGGAGGGGAAGTGCCTTATATACAAAGGTGAAGTGCCTT 363

QY 370 GGCACCAACTTCCCGCGACGCGCCCGTGTATGAAGAACAAAGAGCGCGCTGGAGCCC 429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 GGTACCAATTTTCTGCTGTATGGCCCGTGTATGAAGAACAAATCAGGAGGATGGAGCCA 423

QY 430 AGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGTATGSCCCTG 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 AGCACTGAGTGGTTTATCCAGAGAAATGGTGTCTGTGTGACGTAATGTGTATGSCCCTT 483

QY 490 AAGTGGGCGACCGGCACCTGATCTGCCACCACTACACCAAGTACCGGACCAAGAGGCC 549
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 AAAGTCGGTGATCGTCTGTTGATCTGCCATCACTATATCTTTACAGGTCCAAGAAAGCA 543

QY 550 GTGCGCGCCCTGACCATGCCCGGCTTCCACTTACCGACATCCGGCTCCAGATGTCGCGG 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 GTCCGTGCTTGACAAATGCCAGGATTTCAATTTTACAGACATCCGCCTTCAGATGCTGAGG 603

QY 610 AAGAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGGTACAGCGACCTG 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 AAAGAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTATCTT 663

QY 670 CCCGAGAGGCCCAACTGA 687
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 CCTGAAAAAGCAAAATTGA 681
```

RESULT 11

```

US-09-976-673-23
; Sequence 23, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Fradkov, Arcady
; APPLICANT: Labas, Yulii
; APPLICANT: Matz, Mikhail
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Gurskaya, Nadezda
```

```

; TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
; FILE REFERENCE: CLON-028WO
; CURRENT APPLICATION NUMBER: US/09/976,673
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,018
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60,306,131
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 684
; TYPE: DNA
; ORGANISM: heteractis crispa
US-09-976-673-23

Query Match      63.8%; Score 438.2; DB 9; Length 684;
Best Local Similarity 77.6%; Pred. No. 5.8e-103;
Matches 530; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 5 TGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TGGCTGTTTGTGTAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATG 61

QY 65 GCCACTACTTCAAGTCGAGGGCGAGGGCGACGGCAACCCCTTCGCGCGCACCCAGAGCA 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 GCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATTACAGGTACGCAGAGCA 121

QY 125 TGAGAAATCCACGTGACCGAGGGCGCCCTTCGCTTCGACATCCTGGCCCCCT 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 TGAGGATTATGTCACCGAAGGGCTCCATTACCATTTGCTTCGACATTTTGGCACCGT 181

QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCCGAGATCCCGACTTCTTCA 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATAACGGCAGAGATTCCTGATTTCTTCA 241

QY 245 AGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCAACCACTACGAGGACGGCGGCA 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 AGCAGTCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCACTATGAAGATGGAGGCA 301

QY 305 TCCTGACCGCCACCAAGACACCAAGCTGGAGGGCAACTGCTGATCTACAAGGTGAAGG 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 TTCTTACTGCTCATCAGGACACAAAGCCTGGAGGGGAACCTGCTTATATACAAGGTGAAG 361

QY 365 TGCACGGCACCAACTTCCCCCGCACCGCCCTGATGAAGAAACAAGAGCGCGCTGGG 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TCCTTGGTACCAATTTTCTGCTGATGGCCCCCGTGA TGAAGAAACAATCAGGAGGATGGG 421

QY 425 AGCCACAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCCGGAACGTGATGG 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AGCCAAGCACTGAGGTGGTTTATCCAGAGAAATGGTGTCTGTGTGACGTAATGTGATGG 481

QY 485 CCCTGAAGTGGCGCGACCGGCACCTGATCTGCCACCACTACACCACTACCGGAGCAAGA 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 CCCTTAAAGTCGGTGATCGTCGTTTGTATCTGCCATCTCTATATACTTCTTACAGGTCCAAGA 541

QY 545 AGGCGGTGCGGCCCTTGACCATGCCCGCTTCCACTTCACCGACATCCGGCTCCAGATGC 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAGCAGTCCGTGCTTGACAATGCCAGGATTTTCATTTTACAGACATCCGCCCTTCAGATGC 601

QY 605 TCGGGAAGAAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGGTACAGCG 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 CGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTG 661

QY 665 ACCTGCCCGAGAGGCCCAACTGA 687
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATCTTCTCTGAAAAAGCAAAATTGA 684
```

RESULT 12

```

US-09-976-673-5
; Sequence 5, Application US/09976673
; Patent No. US20020160473A1
```


QY 545 AGCCCGTGGCGCCCTGACCATGCCGGCTTCCACTTCCAGACATCCGGCTCCAGATGC 604
Db 618 AAGCAGTCCGTGCCTTGACAATGCCAGGATTTCAATTTTACAGACATCCGCCCTTCAGATGC 677
QY 605 TGCGGAAAGAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664
Db 678 CGAGGAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTG 737
QY 665 ACCTGCCCGAGAAGGCCAACTGA 687
Db 738 ATCTTCTGAAAAAGCAAATTGA 760

RESULT 14
US-10-155-809-15
; Sequence 15, Application US/10155809
; Publication No. US20030092884A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Chudakov, Dmitzy
; TITLE OF INVENTION: Kindling Fluorescent Proteins and
; FILE OF INVENTION: Methods for Their Use
; FILE REFERENCE: CLON-073
; CURRENT APPLICATION NUMBER: US/10/155,809
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,752
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/329,176
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/210,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/976,673
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Anthozoa

US-10-155-809-15

Query Match 63.6%; Score 436.6; DB 14; Length 760;
Best Local Similarity 77.5%; Pred. No. 1.5e-102;
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 5 TGAGCGGCGCTGCTGAAGGAGAGTATGCGGCATCAAGATGTACATGGAGGGCACCGTGAACG 64
Db 78 TGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATG 137
QY 65 GCCACTACTTCAAGTGGAGGGCGAGGGCGACGGCAACCCCTTCGCCCTTCGACATCCTGGCCCCCT 124
Db 138 GCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATTTCAGGTACGCAGAGCA 197
QY 125 TGAGATCCACGTGACCGAGGGCGCCCCCTGCCCCCTTCGCCCTTCGACATCCTGGCCCCCT 184
Db 198 TGAGGATTTCATGTACCCGAAGGGGCTCCATTACCATTTGGCTTCGACATTTTGGCACCGT 257
QY 185 GTCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCGGAGATCCCCGACTTCTTCA 244
Db 258 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTTCCCGATTCTTCA 317
QY 245 AGCAGAGTTCCCCGAGGGCTTCACTTGGGAGAGAACCACCACTACGAGGACGCGGGCA 304
Db 318 AGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGAGGCA 377
QY 305 TCCTGACCGCCCCACAGGACACAGCCCTGGAGGGCAACTGCCTGTACTACAAAGTGAAGG 364
Db 378 TTCTTACTGCTCATCAGGACACAAGCCTGGAGGGAACTGCCTTATATACAAGTGAAG 437

QY 365 TGACCGGCACCAACTTCCCGCGCGAGCGCCCGTGTATGAAGAACAAGAGCGGCGGTGGG 424
Db 438 TCCTTGGTACCAATTTTCTGTGTATGGCCCCGTGTATGAAGAACAATAATCAGGAGGATGGG 497
QY 425 AGCCAGACCGAGGTGGTGTACCCCGAGAAACGGCGTGTGTGGGCCGGAACGTGATGG 484
Db 498 AGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTCTCTGTGGACGTAGCGTGATGG 557
QY 485 CCCTGAAGGTGGGCGACCGGCACCTGTATCTGCCACCACTACACCAGCTACCGGAGCAAGA 544
Db 558 CCCTTAAAGTCGGTGTATCGTCTGTTGATCTGCCATCACTATACTTCTTACAGGTCCAAGA 617
QY 545 AGGCGGTGCGGCCCTGACCATGCGCGGCTTCCACTTACCGACATCCGGCTCCAGATGC 604
Db 618 AAGCAGTCCGTGCCTTGACAATGCCAGGATTTCAATTTTACAGACCACCGCCTTCAGATGC 677
QY 605 TGCGGAAAGAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGTACAGCG 664
Db 678 GGAGGAAAGAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTG 737
QY 665 ACCTGCCCGAGAAGGCCAACTGA 687
Db 738 ATCTTCTGAAAAAGCAAATTGA 760

RESULT 15
US-09-976-673-1
; Sequence 1, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Fradkov, Arcady
; APPLICANT: Labas, Yulii
; APPLICANT: Matz, Mikhail
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Gurskaya, Nadezda
; TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
; FILE REFERENCE: CLON-028WO
; CURRENT APPLICATION NUMBER: US/09/976,673
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,018
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60,306,131
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 910
; TYPE: DNA
; ORGANISM: heteractis crispa
US-09-976-673-1

Query Match 63.3%; Score 435; DB 9; Length 910;
Best Local Similarity 77.3%; Pred. No. 3.9e-102;
Matches 528; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 5 TGAGCGGCGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACG 64
Db 83 TGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATG 142
QY 65 GCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCCGCACCCAGAGCA 124
Db 143 GCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATTACAGGTACGCAGAGCA 202
QY 125 TGAGAATCCACGTGACCGAGGGCGCCCCCTGCCCCCTTCGCCCTTCGACATCCTGGCCCCCT 184
Db 203 TGAGGATTTCATGTACCCGAAGGGGCTCCATTACCATTTGGCTTCGACATTTTGGCACCGT 262
QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCGGAGATCCCCGACTTCTTCA 244
Db 263 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTTCCCGATTCTTCA 322
QY 245 AGCAGAGCTTCCCCGAGGGCTTCACTTGGGAGAGAACCAACCACTACGAGGACGCGGGCA 304

| | | | |
|----|-----|--|-----|
| Db | 323 | AGCAGTCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGGAGGCA | 382 |
| Qy | 305 | TCCTGACCGCCCAACAGACACAGCCCTGGAGGGCAACTGCCCTGATCTACAAGGTGAAGG | 364 |
| Db | 383 | TTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAACTGCCCTTATATACAAGGTGAAG | 442 |
| Qy | 365 | TGCACGGCACCAACTTCCCCCGCCGACGGCCCCCGTGATGAAGAAACAAGACGGCGGCTGGG | 424 |
| Db | 443 | TCCTTGGTACCAATTTTCCTGCTGATGGCCCCCGTGATGAAGAAACAATCAGGAGGATGGG | 502 |
| Qy | 425 | AGCCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGCTGTGCGGCCCGGAACGTGATGG | 484 |
| Db | 503 | AGCCATGCACTGAGGTGGTGTATCCAGAGATGGTGTCTGTGTGGACGTAATGTGATGG | 562 |
| Qy | 485 | CCCTGAAGGTGGCGGACCGGCACCTGATCTGCCACCACTACACCCAGCTACCGGAGCAAGA | 544 |
| Db | 563 | CCCTTAAAGTCGGTGATCGTCTGTTGATCTGCCATCTCTATACTTCTTACAGGTCCAAGA | 622 |
| Qy | 545 | AGGCCGTGCGCCCTGACCATGCCCGGCTTCCACTTCACCGACATCCGGCTCCAGATGC | 604 |
| Db | 623 | AAGCAGTCCGTGCCCTTGACAAATGCCAGGATTTCAATTTTACAGACATCCGCCCTTCAGATGC | 682 |
| Qy | 605 | TGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG | 664 |
| Db | 683 | CGAGGAAAACGAAAGACGAGTACTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG | 742 |
| Qy | 665 | ACCTGCCCCGAGAAGGCCAACTGA | 687 |
| Db | 743 | ATCTTCTGAAAAAGCAATGA | 765 |

Search completed: December 10, 2004, 21:28:39
Job time : 605 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2004, 16:17:09 ; Search time 3701 Seconds
(without alignments)
974.426 Million cell updates/sec

Title: US-09-976-673-11
Perfect score: 687
Sequence: 1 atggtgagcgcctgctgaa.....tgcccgagaaggccaactga 687

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|-------|---------|--------------|-------------|--------------------|
| 1 | 687 | 100.0 | 687 | 6 ABL41172 | ABL41172 H. crispa |
| 2 | 671 | 97.7 | 687 | 6 ABL41173 | ABL41173 H. crispa |
| 3 | 671 | 97.7 | 1396 | 6 ABL41174 | ABL41174 Cr-449-ta |
| 4 | 671 | 97.7 | 1396 | 10 ACA61024 | ACA61024 DNA encod |
| 5 | 669.4 | 97.4 | 1424 | 6 ABL41175 | ABL41175 Cr-449-ta |
| 6 | 669.4 | 97.4 | 1424 | 10 ACA61025 | ACA61025 DNA encod |
| 7 | 475 | 69.1 | 640 | 12 ADO56228 | Ado56228 Red fluor |
| 8 | 446.6 | 65.0 | 1376 | 10 ACA61026 | ACA61026 DNA encod |
| 9 | 441.2 | 64.2 | 681 | 6 ABL41170 | ABL41170 H. crispa |
| 10 | 439.6 | 64.0 | 681 | 6 ABL41171 | ABL41171 H. crispa |
| 11 | 438.2 | 63.8 | 684 | 6 ABL41180 | ABL41180 H. crispa |
| 12 | 436.6 | 63.6 | 684 | 6 ABL41169 | ABL41169 H. crispa |
| 13 | 436.6 | 63.6 | 760 | 10 ABA00802 | Aba00802 Wild type |
| 14 | 436.6 | 63.6 | 760 | 10 ABA00806 | Aba00806 Multiple |
| 15 | 435 | 63.3 | 910 | 6 ABL41167 | ABL41167 H. crispa |
| 16 | 435 | 63.3 | 910 | 6 ABL41182 | ABL41182 H. crispa |
| 17 | 431.8 | 62.9 | 760 | 10 ABA00805 | Aba00805 Multiple |
| 18 | 431.8 | 62.9 | 760 | 10 ABA00804 | Aba00804 Chromopro |
| 19 | 431.8 | 62.9 | 760 | 10 ABA00803 | Aba00803 Chromopro |
| 20 | 431.8 | 62.9 | 908 | 6 ABL41168 | ABL41168 H. crispa |
| 21 | 429.2 | 62.5 | 680 | 6 ABL41181 | ABL41181 H. crispa |

| | | | | | |
|----|-------|------|------|-------------|--------------------|
| 22 | 422.6 | 61.5 | 696 | 3 AAA50885 | Aaa50885 A. sulcat |
| 23 | 417.8 | 60.8 | 707 | 6 AAD46286 | Aad46286 Anemonia |
| 24 | 417.8 | 60.8 | 1398 | 10 ACA61028 | Aca61028 DNA encod |
| 25 | 415.6 | 60.5 | 1404 | 10 ACA61027 | Aca61027 DNA encod |
| 26 | 415.2 | 60.4 | 699 | 6 AAD46287 | Aad46287 Anemonia |
| 27 | 362.6 | 52.8 | 681 | 6 ABA00245 | Aba00245 C. gigant |
| 28 | 361.2 | 52.6 | 835 | 6 ABA00244 | Aba00244 C. gigant |
| 29 | 306 | 44.5 | 678 | 10 ADC24132 | Adc24132 Discosoma |
| 30 | 306 | 44.5 | 678 | 11 ADL46211 | Adl46211 Discosoma |
| 31 | 302.8 | 44.1 | 678 | 11 ADL46282 | Adl46282 Discosoma |
| 32 | 298.8 | 43.5 | 681 | 4 AAH47656 | Aah47656 Anthozoan |
| 33 | 298.2 | 43.4 | 681 | 10 ADC24130 | Adc24130 Discosoma |
| 34 | 298.2 | 43.4 | 681 | 11 ADL46209 | Adl46209 Discosoma |
| 35 | 297.4 | 43.3 | 723 | 8 ABZ22476 | Abz22476 Mammalian |
| 36 | 292.4 | 42.6 | 678 | 6 AAD46278 | Aad46278 Discosoma |
| 37 | 292.4 | 42.6 | 678 | 6 AAD28208 | Aad28208 Discosoma |
| 38 | 292.4 | 42.6 | 678 | 6 AAD28207 | Aad28207 Discosoma |
| 39 | 292.4 | 42.6 | 695 | 3 AAA48743 | Aaa48743 Humanised |
| 40 | 291.8 | 42.5 | 678 | 10 ADC24128 | Adc24128 Discosoma |
| 41 | 291.8 | 42.5 | 678 | 11 ADL46207 | Adl46207 Discosoma |
| 42 | 291.8 | 42.5 | 704 | 11 ADN33980 | Adn33980 Mutant-ty |
| 43 | 291.4 | 42.4 | 681 | 10 ADC24134 | Adc24134 Discosoma |
| 44 | 291.4 | 42.4 | 681 | 11 ADL46205 | Adl46205 Human cod |
| 45 | 291.4 | 42.4 | 681 | 11 ADL46225 | Adl46225 Human cod |

ALIGNMENTS

RESULT 1
ABL41172
ID ABL41172 standard; cDNA; 687 BP.
XX
AC ABL41172;
XX
DT 12-AUG-2002 (first entry)
XX
DE H. crispa fluorescent protein mutant FP10-cr1 encoding cDNA.
XX
KW Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRET;
KW colouring agent; pharmaceutical; cosmetic; immunoassay; gene;
KW fluorescence resonance energy transfer; gene expression; mutant; ss.
XX
OS Heteractis crispa.
XX
FH Key Location/Qualifiers
FT CDS 1..687
FT /tag= a
FT /product= "mutant fluorescent protein"
XX
PN WO200230965-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032080.
XX
PR 12-OCT-2000; 2000US-0240018P.
PR 16-JUL-2001; 2001US-0306131P.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;
XX
DR WPI; 2002-444170/47.
DR P-PSDB; ABB07994.
XX
PT Novel nucleic acid encoding Stichodactylidaen chromoprotein and its
PT fluorescent mutant useful as coloring agent, labels in analyte detection
PT assays, markers in recombinant DNA applications and filters in
PT sunscreens.
XX
PS Claim 5; Fig 10; 81pp; English.
XX

CC The invention relates to a nucleic acid present in other than its natural
CC environment and encoding an Stichodactylidae chromoprotein or its
CC fluorescent mutant, where the fluorescent protein has an emission maximum
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are
CC useful in applications employing a chromo or fluorescent nucleic acid or
CC protein. Recombinant vectors comprising the nucleic acid is useful for
CC producing an Anthozoan chromo and/or fluorescent protein. The
CC chromoproteins, and their fluorescent mutants are useful as colouring
CC agents capable of imparting colour or pigment to a particular composition
CC of matter. The chromoproteins can be incorporated into a variety of
CC different compositions including food compositions, pharmaceuticals,
CC cosmetics, living organisms, e.g. animals and plants, and as labels in
CC analyte detection assays, e.g. assays for biological analytes of interest
CC (see ABL41167 for a detailed description of the various uses of the
CC chromoproteins). The present sequence represents the H. crispa
CC fluorescent protein mutant FP10-cr1 encoding cDNA
XX
SQ Sequence 687 BP; 150 A; 232 C; 212 G; 93 T; 0 U; 0 Other;

Query Match 100.0%; Score 687; DB 6; Length 687;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60
|||
Db 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60
|||
Qy 61 AACGGCCACTACTTCAAGTGGGAGGGGAGGGCGACGGCAACCCCTTCGCGGACCCAG 120
|||
Db 61 AACGGCCACTACTTCAAGTGGGAGGGGAGGGCGACGGCAACCCCTTCGCGGACCCAG 120
|||
Qy 121 AGCATGAGAAATCCAGTACCGAGGGGCGCCCTTCGCTTTCGCTTTCGATCCTGGCC 180
|||
Db 121 AGCATGAGAAATCCAGTACCGAGGGGCGCCCTTCGCTTTCGATCCTGGCC 180
|||
Qy 181 CCTGCTGCGAGTACGGGAGGAGGACCTTCGTGACCAACACCGCGGAGTCCCGACTTC 240
|||
Db 181 CCTGCTGCGAGTACGGGAGGAGGACCTTCGTGACCAACACCGCGGAGTCCCGACTTC 240
|||
Qy 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCCACCTACGAGGACGGC 300
|||
Db 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCCACCTACGAGGACGGC 300
|||
Qy 301 GGATCCTGACCGCCACCAAGGACACCACTGGAGGGCAACTGCTGATCTACAAGGTG 360
|||
Db 301 GGATCCTGACCGCCACCAAGGACACCACTGGAGGGCAACTGCTGATCTACAAGGTG 360
|||
Qy 361 AAGTGACGCGCACCACTTCCCGCGGACCGGCGGCTGATGAAGAACAGAGCGGCGG 420
|||
Db 361 AAGTGACGCGCACCACTTCCCGCGGACCGGCGGCTGATGAAGAACAGAGCGGCGG 420
|||
Qy 421 TGGAGCCGAGCAGCGAGGTGGTGTACCCCGAGAACGGCGGTGCTGCGGCGGAAACGTG 480
|||
Db 421 TGGAGCCGAGCAGCGAGGTGGTGTACCCCGAGAACGGCGGTGCTGCGGCGGAAACGTG 480
|||
Qy 481 ATGGCCCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACCACTACCGGAGC 540
|||
Db 481 ATGGCCCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACCACTACCGGAGC 540
|||
Qy 541 AAGAAGCCGTCGCGCCCTGACCATGCCCCGCTTCCACTTCACCGACATCCGCTCCAG 600
|||
Db 541 AAGAAGCCGTCGCGCCCTGACCATGCCCCGCTTCCACTTCACCGACATCCGCTCCAG 600
|||
Qy 601 ATGCTGCGGAAGAAGAGGAGTACTTTCGAGTGTACGAGGCGGCGGCTGCGGCTAC 660
|||
Db 601 ATGCTGCGGAAGAAGAGGAGTACTTTCGAGTGTACGAGGCGGCGGCTGCGGCTAC 660
|||
Qy 661 ACCGACCTGCCCGAGAGGCGCAACTGA 687
|||
Db 661 ACCGACCTGCCCGAGAGGCGCAACTGA 687
|||

ABL41173
ID ABL41173 standard; cDNA; 687 BP.
XX
AC ABL41173;
XX
DT 12-AUG-2002 (first entry)
XX
DE H. crispa alternative fluorescent protein mutant FP10-cr1 cDNA.
XX
KW Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW fluorescence resonance energy transfer; gene expression; mutant; ss.
XX
OS Heteractis crispa.
XX
FH Key Location/Qualifiers
FT CDS 1..687
FT /tag= a
FT /product= "mutant fluorescent protein"
FT misc_feature 4..6
FT /tag= b
FT /note= "the amino acid Val encoded by the above codon is
FT not indicated in the corresponding protein"
XX
WO200230965-A2.
PN
XX
PD 18-APR-2002.
XX
PX 12-OCT-2001; 2001WO-US032080.
XX
PF 12-OCT-2000; 2000US-0240018P.
PR 16-JUL-2001; 2001US-0306131P.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;
XX
DR WPI; 2002-444170/47.
DR P-PSDB; ABB07995.
XX
PT Novel nucleic acid encoding Stichodactylidae chromoprotein and its
PT fluorescent mutant useful as coloring agent, labels in analyte detection
PT assays, markers in recombinant DNA applications and filters in
PT sunscreens.
XX
PS Claim 5; Fig 10; 81pp; English.
XX
CC The invention relates to a nucleic acid present in other than its natural
CC environment and encoding an Stichodactylidae chromoprotein or its
CC fluorescent mutant, where the fluorescent protein has an emission maximum
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are
CC useful in applications employing a chromo or fluorescent nucleic acid or
CC protein. Recombinant vectors comprising the nucleic acid is useful for
CC producing an Anthozoan chromo and/or fluorescent protein. The
CC chromoproteins, and their fluorescent mutants are useful as colouring
CC agents capable of imparting colour or pigment to a particular composition
CC of matter. The chromoproteins can be incorporated into a variety of
CC different compositions including food compositions, pharmaceuticals,
CC cosmetics, living organisms, e.g. animals and plants, and as labels in
CC analyte detection assays, e.g. assays for biological analytes of interest
CC (see ABL41167 for a detailed description of the various uses of the
CC chromoproteins). The present sequence represents the H. crispa
CC alternative fluorescent protein mutant FP10-cr1 encoding cDNA
XX
SQ Sequence 687 BP; 145 A; 230 C; 219 G; 93 T; 0 U; 0 Other;

Query Match 97.7%; Score 671; DB 6; Length 687;
Best Local Similarity 98.5%; Pred. No. 8e-109;
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60
|||
Db 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60
|||

QY 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGACCCAG 120
Db 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGACCCAG 120
QY 121 AGCATGAGATCCACGTACCGAGGGCGCCCTTCGCCCTTCGACATCTGGCC 180
Db 121 AGCATGGGATCCACGTACCGAGGGCGCCCTTCGCCCTTCGACATCTGGCC 180
QY 181 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTC 240
Db 181 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTC 240
QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCTACGAGACGGC 300
Db 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCTACGAGACGGC 300
QY 301 GGATCTCTGACCGCCACACGAGACACGACCTGGAGGGCACTGCCTGATCTACAAGGTG 360
Db 301 GGATCTCTGACCGCCACACGAGACACGACCTGGAGGGCACTGCCTGATCTACAAGGTG 360
QY 361 AAGGTGCACGGCACCAACTTCCCGCGGACGGCCCTGTGATGAAGAACAGAGCGCGGC 420
Db 361 AAGGTGTGGGCACCAACTTCCCGCGGACGGCCCTGTGATGAAGAACAGAGCGCGGC 420
QY 421 TGGGAGCCAGCACCGAGGTGGTGTACCCCGAGAACGGGGTCTGTGGCGCCGGAACGTG 480
Db 421 TGGGAGCCAGCACCGAGGTGGTGTACCCCGAGAACGGGGTCTGTGGCGCCGGAACGTG 480
QY 481 ATGGCCCTGAAGGTGGCGACCGGCACCTGATCTGCCACCACTACACGACTACCGGAGC 540
Db 481 ATGGCCCTGAAGGTGGCGACCGGCAGCCGGCGGTGATCTGCCACCACTACACGAGCTACCGGAGC 540
QY 541 AAGAGGCGGTGCGCGCCCTGACCATGCCCGGCTTCCACTTACCGACATCCGGTCCAG 600
Db 541 AAGAGGCGGTGCGCGCCCTGACCATGCCCGGCTTCCACTTACCGACATCCGGTCCAG 600
QY 601 ATGCTGCGGAAGAGAGGACGAGTACTTTCGAGCTGTACAGGCCAGCGTGGCCCGGTAC 660
Db 601 ATGCTGCGGAAGAGAGGACGAGTACTTTCGAGCTGTACAGGCCAGCGTGGCCCGGTAC 660
QY 661 AGCGACTGCCCGGAGAGGCGCACTGA 687
Db 661 AGCGACTGCCCGGAGAGGCGCACTGA 687

RESULT 3
ABL41174

ID ABL41174 standard; DNA; 1396 BP.

AC ABL41174;

DT 12-AUG-2002 (first entry)

DE Cr-449-tandem fusion protein nucleotide sequence.

KW Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRET;
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW fluorescence resonance energy transfer; fusion protein; ds.

OS Synthetic.

OS Heteractis crispa.

XX Key Location/Qualifiers
FT CDS 14..1396

FT /*tag= a

FT /product= "Cr-449-tandem fusion protein"

FT misc_feature 698..709

FT /*tag= b

FT /note= "linker nucleotide fragment"

PN WO200230965-A2.

XX

PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032080.

PF

XX 12-OCT-2000; 2000US-0240018P.

PR 16-JUL-2001; 2001US-0306131P.

XX

PA (CLON-) CLONTECH LAB INC.

XX

PI Lukyanov SA, Frackov AF, Lukyanov KA, Gurskaya NG;

XX

DR WPI; 2002-444170/47.

DR P-PSDB; ABB07996.

XX

PT Novel nucleic acid encoding Stichodactylidaen chromoprotein and its
PT fluorescent mutant useful as coloring agent, labels in analyte detection
PT assays, markers in recombinant DNA applications and filters in
PT sunscreens.

XX

PS Disclosure; Fig 12; 81pp; English.

XX

CC The invention relates to a nucleic acid present in other than its natural
CC environment and encoding an Stichodactylidaen chromoprotein or its
CC fluorescent mutant, where the fluorescent protein has an emission maximum
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are
CC useful in applications employing a chromo or fluorescent nucleic acid or
CC protein. Recombinant vectors comprising the nucleic acid is useful for
CC producing an Anthozoan chromo and/or fluorescent protein. The
CC chromoproteins, and their fluorescent mutants are useful as colouring
CC agents capable of imparting colour or pigment to a particular composition
CC of matter. The chromoproteins can be incorporated into a variety of
CC different compositions including food compositions, pharmaceuticals,
CC cosmetics, living organisms, e.g. animals and plants, and as labels in
CC analyte detection assays, e.g. assays for biological analytes of interest
CC (see ABL41167 for a detailed description of the various uses of the
CC chromoproteins). The present sequence represents the H. crispa Cr-449-
CC tandem fusion protein nucleotide sequence

XX Sequence 1396 BP; 293 A; 471 C; 444 G; 188 T; 0 U; 0 Other;

Query Match 97.7%; Score 671; DB 6; Length 1396;

Best Local Similarity 98.5%; Pred. No. 7.9e-109;

Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGTGACGGCCCTGCTGAAGGAGAGTATGCCCATCAAGATGTACATGGAGGCCCGTG 60

Db 710 ATGGTGACGGCCCTGCTGAAGGAGAGCATGGCATCAAGATGTACATGGAGGCCCGTG 769

QY 61 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGACCCAG 120

Db 770 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGACCCAG 829

QY 121 AGCATGAGATCCACGTACCGAGGGCGCCCTTCGCCCTTCGACATCCTGGCC 180

Db 830 AGCATGGGATCCACGTACCGAGGGCGCCCTTCGCCCTTCGACATCCTGGCC 889

QY 181 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCCGACTTC 240

Db 890 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCCGACTTC 949

QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCTACGAGGACGGC 300

Db 950 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCTACGAGGACGGC 1009

QY 301 GGATCTCTGACCGCCACACGAGACACGACCTGGAGGGCAACTGCCTGATCTACAAGGTG 360

Db 1010 GGATCTCTGACCGCCACACGAGACACGACCTGGAGGGCAACTGCCTGATCTACAAGGTG 1069

QY 361 AAGGTGCACGGCACCAACTTCCCGCGGACGGCCCGCTGATGAAGAACCAAGAGCGCGGC 420

Db 1070 AAGGTGTCGGGCACCAACTTCCCGCGGACGGCCCGCTGATGAAGAACCAAGAGCGCGGC 1129

QY 421 TGGGAGCCAGCAGGAGGTGTGTACCCCGAGAACGGCGTGTGTGCGCGCGGAACGTG 480

Db 1130 TGGAGCCAGCACCGAGTGGTGTACCCCGAGAACGGCGTGTGTGCGCCGGAACGTG 1189
Qy 481 ATGGCCCTGAAGTGGGCGGACCGGACCTGATGTGCCACCACTACACCAAGCTACCGGAGC 540
Db 1190 ATGGCCCTGAAGTGGGCGGACCGGCGGTGATCTGCCACCACTACACCAAGCTACCGGAGC 1249
Qy 541 AAGAAGGCCGTGCGGCCCTGACCATGCCCCGGCTTCCACTTACCCGACATCCCGGCTCCAG 600
Db 1250 AAGAAGGCCGTGCGGCCCTGACCATGCCCCGGCTTCCACTTACCCGACATCCCGGCTCCAG 1309
Qy 601 ATGCTGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db 1310 ATGCTGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 1369
Qy 661 AGCGACCTGCCCGAGAGGCCAACTGA 687
Db 1370 AGCGACCTGCCCGAGAGGCCAACTGA 1396

RESULT 4
ACA61024
ID ACA61024 standard; DNA; 1396 BP.
XX
AC ACA61024;
XX
DT 09-JUL-2003 (first entry)
XX
DE DNA encoding chromoprotein Cr-449-tandem.
XX
KW Chromo/fluorescent domain; labeled fusion protein;
KW site-specific gene modification; chromoprotein; coloring agent;
KW food composition; pharmaceutical; cosmetic; Cr-449-tandem; gene; ds.
OS Anthozoa.

XX
XX
FH Key
CDS Location/Qualifiers
14..1396
/*tag= a
/product= "Cr-449-tandem"

WO2003031590-A2.
XX
PD 17-APR-2003.
XX
PF 10-OCT-2002; 2002WO-US032560.
XX
PR 12-OCT-2001; 2001US-00976673.
PR 11-FEB-2002; 2002US-0356225P.
PR 22-MAY-2002; 2002US-0383336P.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA;
XX
DR WPI; 2003-381709/36.
DR P-PSDB; ABU09921.
XX

PT New nucleic acid encoding polypeptide products having at least two linked
PT chromo/fluorescent domains, useful for generating transgenic plants or
PT animals or site-specific gene modifications in cell lines.
XX
PS Disclosure; Fig 1; 68pp; English.
XX
CC The invention describes a nucleic acid encoding a polypeptide product
CC comprising a first and a second chromo/fluorescent domain, optionally
CC joined by a linking domain. The first and second chromo/fluorescent
CC domains associate with each other under intracellular conditions so that
CC the encoded polypeptide assumes a tertiary structure. The nucleic acid
CC and the protein are useful in producing labeled fusion proteins that have
CC a precise and predictable signal to fusion partner ratio. The nucleic
CC acid may also be used in generating transgenic, non-human plants or
CC animals or site-specific gene modifications in cell lines. The

CC Chromoproteins may be used as colouring agents, as a food composition, in
CC pharmaceuticals or cosmetics, as labels in analyte detection assays or as
CC selectable markers in recombinant DNA applications. This sequence encodes
CC chromo/fluorescent domain fusion protein Cr-449-tandem
XX
SQ Sequence 1396 BP; 293 A; 471 C; 444 G; 188 T; 0 U; 0 Other;

Query Match 97.7%; Score 671; DB 10; Length 1396;
Best Local Similarity 98.5%; Pred. No. 7.9e-109;
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATGCTGAGCGGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60
Db 710 ATGCTGAGCGGCTGCTGAAGGAGAGCATGCGCATCAAGATGTACATGAGGGCACCGTG 769
Qy 61 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120
Db 770 AACGGCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 829
Qy 121 AGCATGAGAATCCACGTGACCGAGGGCGGCCCTTCCCTTCCCTTCCGATCCTGGCC 180
Db 830 AGCATGCGGATCCACGTGACCGAGGGCGGCCCTTCCCTTCCGATCCTGGCC 889
Qy 181 CCTGCTGCGAGTACGGCAGCAGGACCTTCTGTCACACCGCCGAGATCCCGACTTC 240
Db 890 CCTGCTGCGAGTACGGCAGCAGGACCTTCTGTCACACCGCCGAGATCCCGACTTC 949
Qy 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAGAGAACCACTACGAGGACGGC 300
Db 950 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAGAGAACCACTACGAGGACGGC 1009
Qy 301 GGCATCCTGACCGCCCAACAGGACACCGCCTTGGAGGGCAACTGCCTGATCTACAAGGTG 360
Db 1010 GGCATCCTGACCGCCCAACAGGACACCGCCTTGGAGGGCAACTGCCTGATCTACAAGGTG 1069
Qy 361 AAGGTGCACGGCACCACTTCCCGAGGGCTTCACTTGGGAGAGAACCACTACGAGGACGGC 420
Db 1070 AAGGTGCTGGGCACCACTTCCCGAGGGCTTCACTTGGGAGAGAACCACTACGAGGACGGC 1129
Qy 421 TGGAGAGCCAGCAGCCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGCCGGAACGTG 480
Db 1130 TGGAGAGCCAGCAGCCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGCCGGAACGTG 1189
Qy 481 ATGCCCCCTGAAGTGGCGACCGGACCTGATCTGCCACCACTACACCAAGCTACCGGAGC 540
Db 1190 ATGCCCCCTGAAGTGGCGACCGGCGGTGATCTGCCACCACTACACCAAGCTACCGGAGC 1249
Qy 541 AAGAAGGCCGTGCGGCCCTGACCATGCCCCGGCTTCCACTTACCCGACATCCCGCTCCAG 600
Db 1250 AAGAAGGCCGTGCGGCCCTGACCATGCCCCGGCTTCCACTTACCCGACATCCCGCTCCAG 1309
Qy 601 ATGCTGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db 1310 ATGCTGCGGAAGAAGAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 1369
Qy 661 AGCGACCTGCCCGAGAGGCCAACTGA 687
Db 1370 AGCGACCTGCCCGAGAGGCCAACTGA 1396

RESULT 5
ABL41175
ID ABL41175 standard; DNA; 1424 BP.
XX
AC ABL41175;
XX
DT 12-AUG-2002 (first entry)
XX
DE Cr-449-tandem-actin fusion protein nucleotide sequence.
XX
KW Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;
KW coloring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW fluorescence resonance energy transfer; fusion protein; ds.

XX (CLON-) CLONTECH LAB INC.
XX PA
XX PI
XX Lukanov SA;
XX WPI; 2003-381709/36.
XX P-PSDB; ABU0922.
XX
PT New nucleic acid encoding polypeptide products having at least two linked
PT chromo/fluorescent domains, useful for generating transgenic plants or
PT animals or site-specific gene modifications in cell lines.
XX
XX
XX Disclosure; Fig 2; 68pp; English.
XX
XX The invention describes a nucleic acid encoding a polypeptide product
XX comprising a first and a second chromo/fluorescent domain, optionally
XX joined by a linking domain. The first and second chromo/fluorescent
XX domains associate with each other under intracellular conditions so that
XX the encoded polypeptide assumes a tertiary structure. The nucleic acid
XX and the protein are useful in producing labeled fusion proteins that have
XX a precise and predictable signal to fusion partner ratio. The nucleic
XX acid may also be used in generating transgenic, non-human plants or
XX animals or site-specific gene modifications in cell lines. The
XX chromoproteins may be used as colouring agents, as a food composition, in
XX pharmaceuticals or cosmetics, as labels in analyte detection assays or as
XX selectable markers in recombinant DNA applications. This sequence encodes
XX chromo/fluorescent domain fusion protein Cr-449-tandem-actin
XX
SQ Sequence 1424 BP; 301 A; 477 C; 452 G; 194 T; 0 U; 0 Other;

Query Match 97.4%; Score 669.4; DB 10; Length 1424;
Best Local Similarity 98.4%; Pred. No. 1.5e-108;
Matches 676; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGTTGAGCGGCTGCTGAAGGAGAGTATCGGCATCAAGATGTACATGGAGGCGCACCGTG 60
Db |||||
14 ATGTTGAGCGGCTGCTGAAGGAGAGTATCGGCATCAAGATGTACATGGAGGCGCACCGTG 73

QY 61 AACGCCCACTACTTCAAGTGGAGGGGCGGAGGGCGGCAACCCCTTCGCGGCGCACCCAG 120
Db |||||
74 AACGCCCACTACTTCAAGTGGAGGGGCGGAGGGCGGCAACCCCTTCGCGGCGCACCCAG 133

QY 121 AGCATGAGAATCCACGTGACCGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db |||||
134 AGCATGCGGATCCACGTGACCGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 193

QY 181 CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGTACACACACCGCGGAGATCCCGGACTTC 240
Db |||||
194 CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGTACACACACCGCGGAGATCCCGGACTTC 253

QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCCACCCTACGAGGACGGC 300
Db |||||
254 TTCAAGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCCACCCTACGAGGACGGC 313

QY 301 GGATCCTGACCGCCCAACAGGACACCGCTGGAGGGCAACTGCCTGATCTACAAGGTG 360
Db |||||
314 GGATCCTGACCGCCCAACAGGACACCGCTGGAGGGCAACTGCCTGATCTACAAGGTG 373

QY 361 AAGTGCACGGCACCAACTTCCCGCGGACGGCCCGGTGATGAAGAACAGAGCGGCGGC 420
Db |||||
374 AAGTGCCTGGGCACCAACTTCCCGCGGACGGCCCGGTGATGAAGAACAGAGCGGCGGC 433

QY 421 TGGAGCCCCAGCACCGAGTGGTGTACCCCGAGAACGGCGTGTGCGGCGGGAACGTG 480
Db |||||
434 TGGAGCCCCAGCACCGAGTGGTGTACCCCGAGAACGGCGTGTGCGGCGGGAACGTG 493

QY 481 ATGGCCCTGAAGGTGGCGACCGGCACCTGATCTGCCACCACTACACAGTACCGGAGC 540
Db |||||
494 ATGGCCCTGAAGGTGGCGACCGGCACCGGCGGTGATCTGCCACCACTACACAGTACCGGAGC 553

QY 541 AAGAAGCCCGTGGCGGCGCTGACCATGCCCCGCTTCCACTTCCACCGACATCCGGCTCCAG 600
Db |||||
554 AAGAAGCCCGTGGCGGCGCTGACCATGCCCCGCTTCCACTTCCACCGACATCCGGCTCCAG 613

QY 601 ATGCTGCGGAAGAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 660
Db |||||
614 ATGCTGCGGAAGAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTAC 673

QY 661 AGCGACCTGCCCGAGAGAGGCCCAACTGA 687
Db |||||
674 AGCGACCTGCCCGAGAGAGGCCCAACAGA 700

RESULT 7

ADO56228
ID ADO56228 standard; DNA; 640 BP.
XX
AC ADO56228;
XX
DT 26-AUG-2004 (first entry)
XX
DE Red fluorescent chromoprotein rGFP polynucleotide.
XX
XX RNA splicing; RNA processing; gene silencing; disease prevention;
KW disease treatment; SpRNAi; splicing-component intron;
KW red fluorescent chromoprotein; rGFP; ds.
XX
OS Synthetic.
XX
PN US2004106566-A1.
XX
PD 03-JUN-2004.
XX
PF 15-MAY-2003; 2003US-00439262.
XX
PR 17-MAY-2002; 2002US-0381651P.
PR 16-SEP-2002; 2002US-0411062P.
PR 12-OCT-2002; 2002US-0418405P.
XX
PA (LINS/) LIN S.
PA (YING/) YING S.
XX
PI Lin S, Ying S;
XX
XX WPI; 2004-419483/39.

XX Inducing RNA splicing/processing-associated gene silencing effects, by
PT constructing recombinant nucleic acid composition, cloning composition
PT into vector, introducing vector into cell, generating RNA transcript of
PT composition.

XX Example 3; SEQ ID NO 12; 33pp; English.

XX The invention describes a method of inducing (M1) RNA splicing/processing
CC -associated gene silencing effects. The method involves: constructing a
CC recombinant nucleic acid composition (I) containing an intron (II)
CC flanked with several exons, where (II) can be cleaved out of the exons by
CC RNA splicing and/or processing for gene silencing, and the exons can be
CC linked together to form a gene with desired function; cloning (I) into an
CC expression-competent vector; introducing the vector into a cell, cells,
CC tissue or in vivo; generating RNA transcript of (I); and releasing the
CC metabolic products of (II) by RNA splicing/processing mechanisms, so as
CC to provide gene silencing effects against the genes containing sequences
CC homologous to (II). (M1) is useful for inducing RNA splicing/processing-
CC associated gene silencing effects and in disease prevention and
CC treatment. This sequence represents a red fluorescent chromoprotein rGFP
CC polynucleotide from which the exons are generated for the creation of an
CC artificial gene for examining the RNA splicing/processing-associated gene
CC silencing effects.

XX Sequence 640 BP; 140 A; 212 C; 199 G; 89 T; 0 U; 0 Other;

Query Match 69.1%; Score 475; DB 12; Length 640;
Best Local Similarity 92.5%; Pred. No. 1.9e-74;
Matches 637; Conservative 0; Mismatches 0; Indels 52; Gaps 11;

QY 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGGGCATCAAGATGTACATGGAGGCGACCGTG 60
Db 1 ATGGTGAGCGGCTGCTGAAGGAGAGTATGGGCATCAAGATGTACATGGAGGCGACCGTG 60
QY 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 120
Db 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCCA- 119
QY 121 AGCATGAGATCCACAGTGACCGAGGGCGGCCCTTCGCCCTTCGACATCCTGGCC 180
Db 120 AGCATGAGATCCACAGTGACCGAGGGCGGCCCTTCGCCCTTCGACATCCTGG-C 178
QY 181 CCCTGCTGCGAGTACGGCAGCAG---GACCTTCGTGCACCAACACCGCCGAGATCCCGGAC 237
Db 179 CCCTGCTGCGAGTACGGCAGCAGCAGCCTTCGTGCACCAACACCGCCGAGATCCCGGA- 237
QY 238 TTCTTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAGAGAACCAACCACTACGAGGAC 297
Db 238 TTCTTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAG--- 276
QY 298 GGCGGCATCCTGACCGCCACCAGGACACAGCCTGGAGGGGCACTGCCTGATCTACAAG 357
Db 277 -----ACCAGGACACAGCCTGGA-GGCAACTGCTGATCTACAAG 316
QY 358 GTGAAGGTGACGGCACCAACTTCCCGCGCAGCGGCCCGTGTATGAAGAACAGAGCGGC 417
Db 317 GTGAAGGTGACGGCACCAACTTCCCGCGCAGCGGCCCG-TGATGAAGAACAGAGCGGC 375
QY 418 GGCTGGGAGCCAGCAGCAGGCTGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAAC 477
Db 376 GGCTGGGAGCCAGCAGCAGGCTGTGTACCCCGAGAAC-GCGTGTGTGCGGCCGGAAC 434
QY 478 GTGATGGCCCTGAAGTGGGCGACCGGCACCTGATCTGCCACCACTACACAGTACCGG 537
Db 435 GTGATGGCCCTGAAGTGGGCGACCGGCACCTGATCTG-CACCACTACACAGTACCGG 493
QY 538 AGCAAGAGCGCGTGCAGCGCCCTGACCATGCCCGCTTCCACTTACCGACATCCGGCTC 597
Db 494 AGCAAGAGCGCGTGCAGCGCCCTGACCATGCCCGCTT-CACTTACCGACATCCGGCTC 552
QY 598 CAGATGCTGCGGAAGAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGG 657
Db 553 CAGATGCTGCGGAAGAGAGGACGAGTACTTCGAGCTG-ACGAGGCCAGCGTGGCCCGG 611
QY 658 TACAGCGACTGCCCGAGAGGCCCACTG 686
Db 612 TACAGCGACTGCCCGAGAGGCCCACTG 640

RESULT 8
ID ACA61026
XX
AC ACA61026;
XX
DT 09-JUL-2003 (first entry)
XX
DE DNA encoding chromoprotein HcRed-Cr1-tandem.
XX
KW Chromo/fluorescent domain; labeled fusion protein;
KW site-specific gene modification; chromoprotein; colouring agent;
KW food composition; pharmaceutical; cosmetic; HcRed-Cr1-tandem; gene; ds.
XX Anthozoa.
OS
FH Key Location/Qualifiers
FT CDS 1..1376
FT /*tag= a
FT /product= "HcRed-Cr1-tandem"
FT /transl_except= (pos:679..680, aa:Asn)
XX
PN WO2003031590-A2.
XX

PD 17-APR-2003.
XX
PF 10-OCT-2002; 2002WO-US032560.
XX
PR 12-OCT-2001; 2001US-00976673.
PR 11-FEB-2002; 2002US-0356225P.
PR 22-MAY-2002; 2002US-0383336P.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA;
XX
DR WPI; 2003-381709/36.
DR P-PSDB; ABU09923.
XX
PT New nucleic acid encoding polypeptide products having at least two linked
PT chromo/fluorescent domains, useful for generating transgenic plants or
PT animals or site-specific gene modifications in cell lines.
XX
PS Disclosure; Fig 3; 68pp; English.
XX
CC The invention describes a nucleic acid encoding a polypeptide product
CC comprising a first and a second chromo/fluorescent domain, optionally
CC joined by a linking domain. The first and second chromo/fluorescent
CC domains associate with each other under intracellular conditions so that
CC the encoded polypeptide assumes a tertiary structure. The nucleic acid
CC and the protein are useful in producing labeled fusion proteins that have
CC a precise and predictable signal to fusion partner ratio. The nucleic
CC acid may also be used in generating transgenic, non-human plants or
CC animals or site-specific gene modifications in cell lines. The
CC chromoproteins may be used as colouring agents, as a food composition, in
CC pharmaceuticals or cosmetics, as labels in analyte detection assays or as
CC selectable markers in recombinant DNA applications. This sequence encodes
CC chromo/fluorescent domain fusion protein HcRed-Cr1-tandem
XX
SQ Sequence 1376 BP; 381 A; 292 C; 357 G; 346 T; 0 U; 0 Other;
Query Match 65.0%; Score 446.6; DB 10; Length 1376;
Best Local Similarity 78.2%; Pred. No. 1.8e-69;
Matches 536; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 3 GGTGAGCGGCTGCTGAAGGAGAGTATCGGCATCAAGATGTACATGGAGGCGACCGTGAA 62
Db 692 GATGCTGCTGTTGTTGAAAGAAAGTATCGGCATCAAGATGTACATGGAGGCGACCGTTAA 751
QY 63 CGGCCACTACTCAAGTGCAGGGCGAGGGCGAGCGCAACCCCTTCGCCGGCACCCAGAG 122
Db 752 TGGCCATTATTCAAGTGTGAAGGAGAGGGAGAGCGGCAACCCATTTCAGGTACGCAGAG 811
QY 123 CATGAGAATCCACGTGACCGAGGGCGGCCCTTCGCCCTTCGACATCCTGGCCCC 182
Db 812 CATGAGGATTCATGTACCGAAGGGGCTCCATTACCATTTGCTTCGACATTTTGGCACC 871
QY 183 CTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCGGAGATCCCCGACTTCTT 242
Db 872 GTGTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCCGATTCTT 931
QY 243 CAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTACGAGGCGCGG 302
Db 932 CAAGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCACTATGAAGATGGAGG 991
QY 303 CATCCTGACCGCCACCAAGGACACCAAGCCTGGAGGGCAACTGCTGTATCTACAAGGTGAA 362
Db 992 CATTCCTTACTGCTCATCAGGACACAAAGCCTGGAGGGGAACTGCTTATATACAAGGTGAA 1051
QY 363 GGTGCACGGCACCAACTTCCCGCGGACGGCCCCCGTGTATGAAGAACAGAGCGCGGCTG 422
Db 1052 AGTCCATGGTACCAATTTTCTCTGTGTATGCCCGCTGTGTATGAAGAACAAATCAGGAGGATG 1111
QY 423 GGAGCCCGACCGAGGTGTTGTATCCCGGAGAACGGCGTGTGTGTGCGGCCGGAACGTGAT 482
Db 1112 GGAGCCCAAGCACTGAGGTGTTGTATCCAGAGAAATGTTGCTCTGTGTGGACGTATGTGAT 1171

| | | |
|--|---|--|
| XX | Heteractis crispa. | |
| OS | | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 1..681 |
| FT | | /*tag= a |
| FT | | /product= "mutant fluorescent protein" |
| XX | | |
| PN | WO200230965-A2. | |
| XX | | |
| PD | 18-APR-2002. | |
| XX | | |
| PF | 12-OCT-2001; 2001WO-US032080. | |
| XX | | |
| PR | 12-OCT-2000; 2000US-0240018P. | |
| PR | 16-JUL-2001; 2001US-0306131P. | |
| XX | | |
| PA | (CLON-) CLONTECH LAB INC. | |
| XX | | |
| PI | Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG; | |
| XX | | |
| DR | WPI; 2002-444170/47. | |
| DR | P-PSDB; ABB07993. | |
| XX | | |
| PT | Novel nucleic acid encoding Stichodactylidaen chromoprotein and its | |
| PT | fluorescent mutant useful as coloring agent, labels in analyte detection | |
| PT | assays, markers in recombinant DNA applications and filters in | |
| PT | sunscreens. | |
| XX | | |
| PS | Claim 5; Fig 8; 81pp; English. | |
| XX | | |
| CC | The invention relates to a nucleic acid present in other than its natural | |
| CC | environment and encoding an Stichodactylidaen chromoprotein or its | |
| CC | fluorescent mutant, where the fluorescent protein has an emission maximum | |
| CC | ranging from 580-660 nm. The polynucleotides and encoded proteins are | |
| CC | useful in applications employing a chromo or fluorescent nucleic acid or | |
| CC | protein. Recombinant vectors comprising the nucleic acid is useful for | |
| CC | producing an Anthozoan chromo and/or fluorescent protein. The | |
| CC | chromoproteins, and their fluorescent mutants are useful as coloring | |
| CC | agents capable of imparting colour or pigment to a particular composition | |
| CC | of matter. The chromoproteins can be incorporated into a variety of | |
| CC | different compositions including food compositions, pharmaceuticals, | |
| CC | cosmetics, living organisms, e.g. animals and plants, and as labels in | |
| CC | analyte detection assays, e.g. assays for biological analytes of interest | |
| CC | (see ABL41167 for a detailed description of the various uses of the | |
| CC | chromoproteins). The present sequence represents the H. crispa | |
| CC | fluorescent protein mutant 44-6 encoding cDNA | |
| XX | | |
| SQ | Sequence 681 BP; 185 A; 145 C; 178 G; 173 T; 0 U; 0 Other; | |
| Query Match 64.0%; Score 439.6; DB 6; Length 681; | | |
| Best Local Similarity 78.0%; Pred. No. 3e-68; | | |
| Matches 529; Conservative 0; Mismatches 149; Indels 0; Gaps 0; | | |
| QY | 10 GGCCTGCTGAAGGAGAGTATGCCATCAAGATGTACATGGAGGGCACCGTGACGGCCAC | 69 |
| Db | | |
| QY | 70 TACTTCAAGTCGAGGGCGGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGA | 129 |
| Db | | |
| QY | 130 ATCCACGTACCGAGGGCGCCCTGCCCTTCGCCTTCGACATCCTGGCCCCCTGCTGC | 189 |
| Db | | |
| QY | 190 GAGTACGGCAGCAGGACCTTCGTGCACCAACCCGCCAGATCCCGACTTCTTCAAGCAG | 249 |
| Db | | |
| QY | 250 AGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGGACGGCGGCATCCTG | 309 |
| Db | | |
| QY | 244 TCTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGGAGGCATCTT | 303 |
| Db | | |
| QY | 310 ACCGCCACAGGACACACAGCCCTGGAGGGCAACTGCCCTGATCTACAAGGTGAAGGTGCAC | 369 |
| Db | | |
| QY | 370 GGCACCAACTTCCCGCCGACGGCCCGTGTATGAAGAACAGAGCGCGGTGGAGGCC | 429 |
| Db | | |
| QY | 430 AGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGATGGCCCTG | 489 |
| Db | | |
| QY | 490 AAGGTGGGACCGGCACCTGATCTGCCACCACTACACCACTACCGAGCAAGAGGCC | 549 |
| Db | | |
| QY | 550 GTGCGGCCCTGACCATGCCCGGCTTCCACTTACCCGACATCCGGCTCCAGATGCTGCGG | 609 |
| Db | | |
| QY | 610 AAGAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGCGGCCCGGTACAGCGACCTG | 669 |
| Db | | |
| QY | 670 CCCGAGAGGCCCAACTGA | 687 |
| Db | | |
| QY | 664 CCTGAAAAGCAAAATTGA | 681 |
| Db | | |
| RESULT 11 | | |
| ABL41180 | ABL41180 standard; cDNA; 684 BP. | |
| ID | ABL41180; | |
| AC | 12-AUG-2002 (first entry) | |
| XX | H. crispa chromoprotein mutant hcCP mut C148S encoding cDNA. | |
| DE | Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRET; | |
| XX | colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene; | |
| KW | fluorescence resonance energy transfer; gene expression; mutant; ss. | |
| KW | Heteractis crispa. | |
| XX | | |
| OS | | |
| XX | | |
| FT | Key | Location/Qualifiers |
| FT | CDS | 1..684 |
| FT | | /*tag= a |
| FT | | /product= "mutant fluorescent protein" |
| XX | | |
| PN | WO200230965-A2. | |
| XX | | |
| PD | 18-APR-2002. | |
| XX | | |
| PF | 12-OCT-2001; 2001WO-US032080. | |
| XX | | |
| PR | 12-OCT-2000; 2000US-0240018P. | |
| PR | 16-JUL-2001; 2001US-0306131P. | |
| XX | | |
| PA | (CLON-) CLONTECH LAB INC. | |
| XX | | |
| PI | Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG; | |
| XX | | |
| DR | WPI; 2002-444170/47. | |
| DR | P-PSDB; ABB07998. | |
| XX | | |
| PT | Novel nucleic acid encoding Stichodactylidaen chromoprotein and its | |
| PT | fluorescent mutant useful as coloring agent, labels in analyte detection | |
| PT | assays, markers in recombinant DNA applications and filters in | |
| PT | sunscreens. | |
| XX | | |
| PS | Example; Fig 15; 81pp; English. | |

Db 62 GCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATTTACAGGTACCGAGAGCA 121
QY 125 TGAGAAATCCACGTGACCGAGGGCCCCCCTGCCCCCTTCGACATCTGGCCCCCCT 184
Db 122 TGAGGATTATGTCACCGAAGGGGCTCCATTACCAATTGCGCTTCGACATTTTGGCACCGT 181
QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCCGACTTCTTCA 244
Db 182 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTTCCCGATTCTTCA 241
QY 245 AGCAGAGCTTCCCCGAGGCTTCACCTGGGAGAGAACACCACTACGAGCGCGGGCA 304
Db 242 AGCAGTCTTTCCCTGAAGGCTTTACTTTGGGAAAGAACCAACCTATGAAGATGGAGGCA 301
QY 305 TCCTGACCGCCACACAGGACACAGCCCTGGAGGGCAACTGCCTGATCTACAGGTGAAGG 364
Db 302 TTCCTACTGCTCATCAGGACACAAGCCCTGGAGGGAACTGCCTTATATACAAAGTGAAG 361
QY 365 TGCACGGCACCAACTTCCCCGGGACGGCCCCCGTGATGAAGAACAAAGAGCGCGCTGGG 424
Db 362 TCCTTGGTACCAATTTTCTGCTGATGGCCCCCGTGATGAAGAACAAATCAGGAGGATGG 421
QY 425 AGCCAGCACCGAGGTGGTGATCCCCCGAGAACGGCGTGCTGTGCGCGGAACTGATGG 484
Db 422 AGCCAAGCACTGAGGTGGTTTATCCAGAGAAATGGTGCTGTGGACGTAATGTGATGG 481
QY 485 CCCTGAAGGTGGGCGACCGGCACCTGATCTGCCACCACTACACCACTACCGGAGCAAGA 544
Db 482 CCCTTAAAGTCGGTGATCGTCTTTGATCTGCCATCTCTATACTTCTTACAGGTCCAAGA 541
QY 545 AGGCGGTGCGGCCCTGACCATGCCCGGCTTCCACTTCACCGACATCCCGGCTCCAGATGC 604
Db 542 AAGCAGTCCGTGCTTGACATGCCAGGATTTTCATTTTACAGACATCCGCGCTCAGATGC 601
QY 605 TGGGAAGAAGAAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGTTACAGCG 664
Db 602 CGAGGAAACGAAAGACGAGTACTTTGAACTGTACGAGCATCTGTGGCTAGGTACAGTG 661
QY 665 ACCTGCCCGAGAGGCCCACTGA 687
Db 662 ATCTTCTGTAAAGCAAAATTGA 684

RESULT 13

ABAA0802
ID ABA00802 standard; cDNA; 760 BP.
XX
AC ABA00802;
XX
DT 01-APR-2003 (first entry)
XX
DE Wild type chromoprotein cDNA.
XX
KW Gene; kindling fluorescent protein; kindling stimulus; movement;
KW labeling; fluorescence resonance energy transfer; FRET;
KW bioluminescence resonance energy transfer; BRET; biosensor;
KW automated screening; ss.
XX
OS Heteractis criepea.
XX
FH Key Location/Qualifiers
CDS 77..760
FT /*tag= a
FT /product= "Chromoprotein"
XX
PN WO200296924-A1.
XX
PD 05-DEC-2002.
XX
PF 24-MAY-2002; 2002WO-US016379.
XX
PR 25-MAY-2001; 2001US-0293752P.
PR 11-OCT-2001; 2001US-0329176P.

XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA, Chudakov D, Lukyanov K;
XX
DR WPI; 2003-156788/15.
DR P-PSDB; AAG79765.
XX
PT Novel nucleic acid that is present in other than its natural environment
PT and that encodes kindling fluorescent protein, is useful in labeling
PT protocols, e.g. labeling proteins, organelles, cells and organisms.
PS Example; Fig 4; 96pp; English.
XX
CC The sequences given in ABA00799-805 encode wild type and mutant kindling
CC fluorescent proteins. The proteins go from a first substantially non-
CC fluorescent or non-fluorescent state to a second fluorescent state upon
CC exposure to a kindling stimulus. The kindling proteins are useful for
CC detecting an entity such as a protein, organelle or cell in a composition
CC such as a cell or a multicellular composition (preferably a multicellular
CC organism), by providing the entity as an entity labeled with the kindling
CC protein, kindling the kindling fluorescent protein label with a kindling
CC stimulus to produce a kindled kindling fluorescent protein label, and
CC exciting the kindled kindling fluorescent protein label with light and
CC detecting any fluorescence from it to detect the entity. The method
CC monitors the movement of the entity. The fluorescent proteins and the
CC cDNA encoding them are useful in labeling protocols, e.g., labeling
CC proteins, organelles, cells and organisms, as biological labels or
CC markers, in protein labeling or tagging applications. The fluorescent
CC kindling proteins are useful as detectable labels, as labels in analyte
CC detection assays, in fluorescence resonance energy transfer (FRET)
CC applications, in bioluminescence resonance energy transfer (BRET)
CC applications, as biosensors in prokaryotic and eukaryotic cells, in
CC applications involving the automated screening of arrays of cells
CC expressing fluorescent reporting groups, in high through-put screening
CC assays, as second messenger detectors, and in fluorescent activated cell
CC sorting assays
XX
SQ Sequence 760 BP; 209 A; 166 C; 189 G; 196 T; 0 U; 0 Other;

Query Match 63.6%; Score 436.6; DB 10; Length 760;
Best Local Similarity 77.5%; Pred. No. 1e-67;
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 5 TGAGCGGCGCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACG 64
Db 78 TGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTAATG 137
QY 65 GCCACTACTTCAAGTCCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA 124
Db 138 GCCATTATTTCAAGTGTGAAGGAGAGGGAGAGACGGCAACCCATTTACAGGTACGCAGAGCA 197
QY 125 TGAGATCCACGTGACCGAGGGCGCCCCCTGCCCCCTTCGCTTCGACATCTTGGCCCCCT 184
Db 198 TGAGGATTATGTCACCGAAGGGGCTCCATTACCAATTGCGCTTCGACATTTTGGCACCGT 257
QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACCAACACCGCGGAGATCCCCGACTTCTTCA 244
Db 258 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCCGATTCTTCA 317
QY 245 AGCAGAGCTTCCCCGAGGGCTTTCACCTGGGAGAGAAACCAACCACTACGAGGACGGCGCA 304
Db 318 AGCAGTCTTTCCCTGAAGGCTTTACTTGGGAAGAAGAACCAACCACTATGAAGATGGAGCA 377
QY 305 TCCTGACCGCCACACAGGACACACAGCCCTGGAGGGCAACTGCCTGATCTACAAGGTGAAG 364
Db 378 TTCCTTACTGCTCATCAGGACACAGCCCTGGAGGGAACTGCCTTATATACAAGGTGAAG 437
QY 365 TGCACGSCACCAACTTCCCCCGCGACGGCCCCCTGATGAAGAACAAAGACGGCGCTGGG 424
Db 438 TCCTTGGTACCAATTTTCTCTGTGATGGCCCCCGTGTATGAAGAACAAATCAGGAGGATGG 497
QY 425 AGCCAGCACCGAGGTGGTGTATCCCCCGAGAACGGCGTGTCTGCGGCCGGAACGCTGATGG 484

Db 498 ACCCATGCACTGAGGTGTTTATCCAGAGATGGTGTCTCTGTGTGACGTAATGTGATGG 557
Qy 485 CCCTGAAGTGGCGACCGGCACCTGATCTGCCACCACTACACCACTACCGGAGCAAGA 544
Db 558 CCCTTAAAGTCGGTGATCGTGGTTGATCTGCCATCTCTATCTTCTTACAGGTCCAAGA 617
Qy 545 AGCCCGTGGCGCCCTGACCATGCCCGGCTTCCACTTACCGACATCCGGCTCCAGATGC 604
Db 618 AAGCAGTCCGTGCTTGACAATGCCAGGATTTCATTTTACAGACATCCGCTTCAGATGC 677
Qy 605 TCCGGAAGAAGAACGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664
Db 678 CGAGGAAAAGAAAGACGAGTACTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG 737
Qy 665 ACCTGCCCGAGAGGCCCAACTGA 687
Db 738 ATCTTCTGAAAAAGCAAAATTGA 760

RESULT 14
ABA00806
ID ABA00806 standard; cDNA; 760 BP.
AC ABA00806;
XX 01-APR-2003 (first entry)

DE Multiple mutant Chromoprotein #2 cDNA.
XX Gene; kindling fluorescent protein; kindling stimulus; movement;
KW labeling; fluorescence resonance energy transfer; FRET;
KW bioluminescence resonance energy transfer; BRET; biosensor;
KW automated screening; ss.

OS Heteractis crispa.
XX Key Location/Qualifiers
FH CDS 77..760
FT /*tag= a
FT /product= "Multiple mutant Chromoprotein #2"
FT replace(182,A)
FT /*tag= c
FT mutation replace(503,T)
FT /*tag= m
FT mutation replace(549..550, AT)
FT /*tag= n
FT mutation replace(594,T)
FT /*tag= h
FT mutation replace(662..663, AT)
FT /*tag= j
FT mutation replace(678,C)
FT /*tag= o
FT mutation replace(686,A)
FT /*tag= i

WO200296924-A1.
XX 05-DEC-2002.
XX 24-MAY-2002; 2002WO-US016379.
XX 25-MAY-2001; 2001US-0293752P.
PR 11-OCT-2001; 2001US-0329176P.
XX (CLON-) CLONTECH LAB INC.

XX Lukyanov SA, Chudakov D, Lukyanov K;
XX WPI; 2003-156788/15.
DR P-PSDB; AAG79769.
XX Novel nucleic acid that is present in other than its natural environment

PT and that encodes kindling fluorescent protein, is useful in labeling
PT protocols, e.g. labeling proteins, organelles, cells and organisms.
XX
PS Example; Fig 8; 96pp; English.
XX
CC The sequences given in ABA00799-805 encode wild type and mutant kindling
CC fluorescent proteins. The proteins go from a first substantially non-
CC fluorescent or non-fluorescent state to a second fluorescent state upon
CC exposure to a kindling stimulus. The kindling proteins are useful for
CC detecting an entity such as a protein, organelle or cell in a composition
CC such as a cell or a multicellular composition (preferably a multicellular
CC organism), by providing the entity as an entity labeled with the kindling
CC protein, kindling the kindling fluorescent protein label with a kindling
CC stimulus to produce a kindled kindling fluorescent protein label, and
CC exciting the kindled kindling fluorescent protein label with light and
CC detecting any fluorescence from it to detect the entity. The method
CC monitors the movement of the entity. The fluorescent proteins and the
CC cDNA encoding them are useful in labeling protocols, e.g., labeling
CC proteins, organelles, cells and organisms, as biological labels or
CC markers, in protein labeling or tagging applications. The fluorescent
CC kindling proteins are useful as detectable labels, as labels in analyte
CC detection assays, in fluorescence resonance energy transfer (FRET)
CC applications, in bioluminescence resonance energy transfer (BRET)
CC applications, as biosensors in prokaryotic and eukaryotic cells, in
CC applications involving the automated screening of arrays of cells
CC expressing fluorescent reporting groups, in high throughput screening
CC assays, as second messenger detectors, and in fluorescent activated cell
CC sorting assays

XX Sequence 760 BP; 208 A; 167 C; 193 G; 192 T; 0 U; 0 Other;
SQ
Query Match 63.6%; Score 436.6; DB 10; Length 760;
Best Local Similarity 77.5%; Pred. No. 1e-67;
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 5 TGAGCGGCTGCTGAAGGAGAGATGTCGCATCAAGATGTACATGGAGGCACCGTGAACG 64
Db 78 TGGCTGGTTTGTGAAAGAAAGTATCGGCATCAAGATGTACATGGAAGGCACGGTTAATG 137
Qy 65 GCCACTACTTCAAGTGCAGGGCGAGGGCGAGGGCGAGCCCAACCCCTTCGCGGCACCCAGAGCA 124
Db 138 GCCATTATTTCAGTGTGAAGGAGAGAGGGAGAGCGCAACCCATTTCGAGGTACGCAGAGCA 197
Qy 125 TGAGATCCACGTGACCGAGGGCGCCCTTCGCTTTCGCTTTCGACATCTGGCCCTT 184
Db 198 TGAGGATTTCATGTCACCGAAGGGCTCCATTACCATTTGCCCTTCGACATTTGGCACCGT 257
Qy 185 GCTCGAGTACGGCAGCAGGACCTTCGTGACCAACACCGCCGAGATCCCGGACTTCTTCA 244
Db 258 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCGCGATTCTTCA 317
Qy 245 AGCAGAGCTTCCCGAGGGCTTCACCTGGAGAGAAACCAACCATCTACGAGGAGCGGCA 304
Db 318 AGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCATCTAGAGATGGAGCA 377
Qy 305 TCCTGACCGCCCAACAGGACACCAAGCTTCGAGGGGCAACTGCCTGATCTACAAGGTGAAG 364
Db 378 TTCCTTACTGCTCATCAGGACACAAAGCTTCGAGGGGAACTGCCTTATATACAAGTGAAG 437
Qy 365 TGCACGGCACCAACTTCCCGCGCAGCGGCCCTGATGAAGAAACAAGAGCGCGCTGGG 424
Db 438 TCCTTGGTACCAATTTCTGCTGATGGCCCGCTGATGAAGAAACAATCAGGAGGATGG 497
Qy 425 AGCCAGCACCGAGGTGTTGATACCCCGAGAACGGCGTGTGTGCGGCGGAAACGTGATGG 484
Db 498 AGCCAAAGCACTGAGGTGTTTATCCAGAGAGATGGTGTCTGTGTGGACGTAGCGTATGG 557
Qy 485 CCCTGAAGGTGGCGACCGGCACCTGATCTGCCACCACTACACCACTACCGGAGCAAGA 544
Db 558 CCCTTAAAGTCGGTGATCGTCTGTTGATCTGCCATCACTATACCTTCTTACAGGTCCAAGA 617
Qy 545 AGGCGGTGCGCGCTGACCATGCGCGCTTCCACTTCACCGACATCCGCTCCAGATGC 604

| | | | |
|----|-----|---|-----|
| Db | 683 | CGAGGAAACGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTG | 742 |
| Qy | 665 | ACCTGCCCGAGAGGCCAACTGA | 687 |
| Db | 743 | ATCTTCCTGAAAAAGCAAAATGA | 765 |

Search completed: December 10, 2004, 19:46:30
Job time : 3703 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2004, 16:17:09 ; Search time 4983 Seconds
(without alignments)
6519.772 Million cell updates/sec

Title: US-09-976-673-11
Perfect score: 687
Sequence: 1 atggtgagcgccctgctgaa.....tgcccgagaaggccaactga 687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| Database : | | | | GenEmbl:* | | | | |
|------------|--|--|--|-----------|----------|--|--|--|
| | | | | 1: | gb_ba:* | | | |
| | | | | 2: | gb_htg:* | | | |
| | | | | 3: | gb_in:* | | | |
| | | | | 4: | gb_om:* | | | |
| | | | | 5: | gb_ov:* | | | |
| | | | | 6: | gb_pat:* | | | |
| | | | | 7: | gb_ph:* | | | |
| | | | | 8: | gb_pl:* | | | |
| | | | | 9: | gb_pr:* | | | |
| | | | | 10: | gb_ro:* | | | |
| | | | | 11: | gb_sts:* | | | |
| | | | | 12: | gb_sy:* | | | |
| | | | | 13: | gb_un:* | | | |
| | | | | 14: | gb_vi:* | | | |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 687 | 100.0 | 687 | 6 | AX527898 Sequence |
| 2 | 671 | 97.7 | 687 | 6 | AX527900 Sequence |
| 3 | 671 | 97.7 | 1396 | 6 | AX527902 Sequence |
| 4 | 669.4 | 97.4 | 1424 | 6 | AX527904 Sequence |
| 5 | 441.2 | 64.2 | 681 | 6 | AX527894 Sequence |
| 6 | 439.6 | 64.0 | 681 | 6 | AX527896 Sequence |
| 7 | 438.2 | 63.8 | 684 | 6 | AX527910 Sequence |
| 8 | 436.6 | 63.6 | 684 | 6 | AX527892 Sequence |
| 9 | 436.6 | 63.6 | 908 | 3 | AF363776 Heteracti |
| 10 | 435 | 63.3 | 910 | 6 | AX527888 Sequence |
| 11 | 435 | 63.3 | 910 | 6 | AX527914 Sequence |
| 12 | 431.8 | 62.9 | 908 | 6 | AX527890 Sequence |
| 13 | 429.2 | 62.5 | 680 | 6 | AX527912 Sequence |
| 14 | 424.8 | 61.8 | 699 | 12 | AY233273 Synthetic |
| 15 | 417.8 | 60.8 | 707 | 6 | AX824739 Sequence |
| 16 | 396.8 | 57.8 | 654 | 6 | AX824741 Sequence |
| 17 | 362.6 | 52.8 | 681 | 6 | AX573118 Sequence |
| 18 | 361.2 | 52.6 | 835 | 3 | AF363775 Condylact |
| 19 | 361.2 | 52.6 | 835 | 6 | AX573116 Sequence |

| | | | | | | | |
|------|-------|------|-------|----|----------|----------|-----------|
| 20 | 358 | 52.1 | 684 | 3 | AF383155 | AF383155 | Condylact |
| 21 | 351.2 | 51.1 | 835 | 3 | AY037777 | AY037777 | Condylact |
| 22 | 306 | 44.5 | 678 | 12 | AF506027 | AF506027 | Synthetic |
| 23 | 299 | 43.5 | 713 | 6 | AX233627 | AX233627 | Sequence |
| 24 | 298.8 | 43.5 | 681 | 6 | AX233584 | AX233584 | Sequence |
| 25 | 298.2 | 43.4 | 681 | 12 | AF506025 | AF506025 | Synthetic |
| 26 | 298.2 | 43.4 | 1395 | 12 | AF506026 | AF506026 | Synthetic |
| 27 | 295 | 42.9 | 10141 | 12 | AY342347 | AY342347 | Red H-Pel |
| 28 | 294 | 42.8 | 10276 | 12 | AY342348 | AY342348 | Red H-Sti |
| 29 | 294 | 42.8 | 10481 | 12 | AY490568 | AY490568 | UAS-Red S |
| 30 | 292.4 | 42.6 | 678 | 6 | AX370404 | AX370404 | Sequence |
| 31 | 292.4 | 42.6 | 678 | 6 | AX370406 | AX370406 | Sequence |
| 32 | 292.4 | 42.6 | 678 | 6 | AX824725 | AX824725 | Sequence |
| 33 | 291.4 | 42.4 | 4692 | 6 | AX463702 | AX463702 | Sequence |
| 34 | 291.4 | 42.4 | 9320 | 6 | AX663075 | AX663075 | Sequence |
| 35 | 290.8 | 42.3 | 6893 | 6 | AX823860 | AX823860 | Sequence |
| C 36 | 289.8 | 42.2 | 8811 | 12 | AY569779 | AY569779 | Cloning v |
| C 37 | 289.8 | 42.2 | 12404 | 12 | AY569780 | AY569780 | Cloning v |
| 38 | 289.2 | 42.1 | 678 | 6 | AX370408 | AX370408 | Sequence |
| 39 | 289.2 | 42.1 | 678 | 6 | AX824732 | AX824732 | Sequence |
| 40 | 288.4 | 42.0 | 678 | 6 | AX824729 | AX824729 | Sequence |
| 41 | 288.2 | 42.0 | 1050 | 6 | AX666133 | AX666133 | Sequence |
| 42 | 288.2 | 42.0 | 6423 | 12 | AY613997 | AY613997 | Cloning v |
| 43 | 287.6 | 41.9 | 675 | 6 | AX824731 | AX824731 | Sequence |
| 44 | 268 | 39.0 | 955 | 3 | AF246709 | AF246709 | Anemonia |
| 45 | 266.4 | 38.8 | 699 | 3 | AF322222 | AF322222 | Anemonia |

ALIGNMENTS

RESULT 1
AX527898
LOCUS AX527898 687 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 11 from Patent WO0230965.
ACCESSION AX527898
VERSION AX527898.1 GI:25172348
KEYWORDS
SOURCE Heteractis crispa (leathery sea anemone)
ORGANISM Heteractis crispa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria; Stichodactylidae; Heteractis.
REFERENCE 1
AUTHORS Lukyanov, S.A., Fradkov, A.F., Lukyanov, K.A. and Gurskaya, N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent: WO 0230965-A 11 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source
1. .687
/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"

| ORIGIN | | Query Match | | 100.0%; Score 687; DB 6; Length 687; | |
|--------|-----|--|-----|--------------------------------------|--|
| | | Best Local Similarity | | 100.0%; Pred. No. 3.5e-79; | |
| | | Matches 687; Conservative | | 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1 | ATGGTGAGCGCCTGCTGAAGGAGAGATATCGGCATCAAGATGTACATGGAGGACCGTG | 60 | | |
| Db | 1 | ATGGTGAGCGCCTGCTGAAGGAGAGATATCGGCATCAAGATGTACATGGAGGACCGTG | 60 | | |
| QY | 61 | AACGGCCACTACTTCAAGTCCGAGGCGAGGGCGACGGCAACCCCTTCGCCGACCCAG | 120 | | |
| Db | 61 | AACGGCCACTACTTCAAGTCCGAGGCGAGGGCGACGGCAACCCCTTCGCCGACCCAG | 120 | | |
| QY | 121 | AGCATGAGATCCACGTGACCGAGGGCGCCCCCTCGCCCTTCGCATCTCGATCTGGCC | 180 | | |
| Db | 121 | AGCATGAGATCCACGTGACCGAGGGCGCCCCCTCGCCCTTCGCATCTCGATCTGGCC | 180 | | |
| QY | 181 | CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCGGAGATCCCCGACTTC | 240 | | |
| Db | 181 | CCCTGCTGCGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCGGAGATCCCCGACTTC | 240 | | |

Db 830 AGCATGCGGATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGACATCCTGGCC 889
QY 181 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCCGAGATCCCCGACTTC 240
Db 890 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCCGAGATCCCCGACTTC 949
QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGC 300
Db 950 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGC 1009
QY 301 GGCACTCTGACCGCCCCACCGAGGACACCGACCTGGAGGGCACTGCCTGATCTACAGGTG 360
Db 1010 GGCACTCTGACCGCCCCACCGAGGACACCGACCTGGAGGGCACTGCCTGATCTACAGGTG 1069
QY 361 AAGGTGACGGCACCAACTTCCCGCCGACGGCCCCCGTGTATGAAGAACAAAGAGCGCGGC 420
Db 1070 AAGGTGTGGGCACCAACTTCCCGCCGACGGCCCCCGTGTATGAAGAACAAAGAGCGCGGC 1129
QY 421 TGGGAGCCAGCACCGAGGTGGTGTATCCCGAGAACCGCGGTGTGTGCGGCCGGAACGTG 480
Db 1130 TGGGAGCCAGCACCGAGGTGGTGTATCCCGAGAACCGCGGTGTGTGCGGCCGGAACGTG 1189
QY 481 ATGGCCCTGAAGGTGGCGGACCGGCACCTGATCTGCCACCACTACACCGACTACCGGAGC 540
Db 1190 ATGGCCCTGAAGGTGGCGGACCGGCAGCTGATCTGCCACCACTACACCGACTACCGGAGC 1249
QY 541 AAGAAGCCCGTGGCGGCCCTGACCATGCCCGCTTCCAATTCAACCGACATCCGGCTCCAG 600
Db 1250 AAGAAGCCCGTGGCGGCCCTGACCATGCCCGCTTCCAATTCAACCGACATCCGGCTCCAG 1309
QY 601 ATGCTGCGGAAGAAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCGGGTAC 660
Db 1310 ATGCTGCGGAAGAGGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCGGGTAC 1369
QY 561 AGCGACTGCCCGAGAGGAGGCCAACTGA 687
Db 1370 AGCGACTGCCCGAGAGGAGGCCAACTGA 1396

RESULT 4
AX527904
LOCUS AX527904 1424 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 17 from Patent WO0230965.
ACCESSION AX527904
VERSION AX527904.1 GI:25172351
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent: WO 0230965-A 17 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source
1 .1424
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="fusion construct"

ORIGIN
Query Match 97.4%; Score 669.4; DB 6; Length 1424;
Best Local Similarity 98.4%; Pred. No. 5.5e-77;
Matches 676; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGGTGAGCGGCCTGCTGAAGGAGAGATATGCGCATCAAGATGTACATGGAGGACCCGTG 60
Db 14 ATGGTGAGCGGCCTGCTGAAGGAGAGCATGCGCATCAAGATGTACATGGAGGACCCGTG 73
QY 61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGACCCAG 120

Db 74 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG 133
QY 121 AGCATGAGAAATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGACATCCTGGCC 180
Db 134 AGCATGCGGATCCACGTGACCGAGGGCGCCCCCTTGCCTTCGACATCCTGGCC 193
QY 181 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCCGAGATCCCCGACTTC 240
Db 194 CCCTGCTGGAGTACGGCAGCAGGACCTTCGTGCACCACACCGCCGAGATCCCCGACTTC 253
QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGC 300
Db 254 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAACCCACCTACGAGACGGC 313
QY 301 GGCACTCTGACCGCCCCACCGAGGACACCGACCTGGAGGGCACTGCCTGATCTACAGGTG 360
Db 314 GGCACTCTGACCGCCCCACCGAGGACACCGACCTGGAGGGCACTGCCTGATCTACAGGTG 373
QY 361 AAGGTGACCGCACCAACTTCCCGCCGACGGCCCCCGTGTATGAAGAACAAAGAGCGCGGC 420
Db 374 AAGGTGTGGGCACCAACTTCCCGCCGACGGCCCCCGTGTATGAAGAACAAAGAGCGCGGC 433
QY 421 TGGGAGCCAGCACCGAGGTGGTGTATCCCGAGAACCGCGGTGTGTGCGGCCGGAACGTG 480
Db 434 TGGGAGCCAGCACCGAGGTGGTGTATCCCGAGAACCGCGGTGTGTGCGGCCGGAACGTG 493
QY 481 ATGGCCCTGAAGGTGGCGGACCGGCACCTGATCTGCCACCACTACACCGACTACCGGAGC 540
Db 494 ATGGCCCTGAAGGTGGCGGACCGCGGCTGATCTGCCACCACTACACCGACTACCGGAGC 553
QY 541 AAGAAGCCCGTGGCGGCCCTGACCATGCCCGCTTCCAATTCAACCGACATCCGGCTCCAG 600
Db 554 AAGAAGCCCGTGGCGGCCCTGACCATGCCCGCTTCCAATTCAACCGACATCCGGCTCCAG 613
QY 601 ATGCTGCGGAAGAAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCGGGTAC 660
Db 614 ATGCTGCGGAAGAGGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCGGGTAC 673
QY 661 AGCGACTGCCCGAGAGGCGCAACTGA 687
Db 674 AGCGACTGCCCGAGAGGCGCAACAGA 700

RESULT 5
AX527894
LOCUS AX527894 681 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 7 from Patent WO0230965.
ACCESSION AX527894
VERSION AX527894.1 GI:25172346
KEYWORDS Heteractis crispa (leathery sea anemone)
SOURCE Heteractis crispa
ORGANISM Heteractis crispa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
Stichodactylidae; Heteractis.
REFERENCE 1
AUTHORS Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent: WO 0230965-A 7 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source
1 .681
/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"

ORIGIN
Query Match 64.2%; Score 441.2; DB 6; Length 681;
Best Local Similarity 78.2%; Pred. No. 1.2e-47;
Matches 530; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
QY 10 GGCCTGCTGAAGGAGAGTATGCCATCAAGATGTACATGGAGGCCACCGTGACGCCAC 69
Db 4 GGTGTTGTAAGAAAGTATGCCATCAAGATGTACATGAAGGCACGGTTAATGGCCAT 63

[illegible]

RESULT 8
AX527892
LOCUS AX527892 684 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 5 from Patent WO230965.
ACCESSION AX527892
VERSION AX527892.1 GI:25172345
KEYWORDS
SOURCE
ORGANISM
Heteractis crispa (leathery sea anemone)
Heteractis crispa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
Stichodactylidae; Heteractis.
1
REFERENCE
AUTHORS Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent.: WO 0230965-A 5 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source Location/Qualifiers
1. .684
/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"

| ORIGIN | | Query Match | Score | DB | Length |
|---------|--------------|-----------------------|------------|--------------------|--------|
| | | Best Local Similarity | 77.5%; | Pred. No. 4.8e-47; | |
| Matches | Conservative | 0; | Mismatches | 154; | Indels |
| 0; | Gaps | 0; | | | |
| QY | 5 | 5 | 5 | 5 | 5 |
| DB | 2 | 2 | 2 | 2 | 2 |
| QY | 65 | 65 | 65 | 65 | 65 |
| DB | 62 | 62 | 62 | 62 | 62 |
| QY | 125 | 125 | 125 | 125 | 125 |
| DB | 122 | 122 | 122 | 122 | 122 |
| QY | 185 | 185 | 185 | 185 | 185 |
| DB | 182 | 182 | 182 | 182 | 182 |
| QY | 245 | 245 | 245 | 245 | 245 |
| DB | 242 | 242 | 242 | 242 | 242 |
| QY | 305 | 305 | 305 | 305 | 305 |
| DB | 302 | 302 | 302 | 302 | 302 |
| QY | 365 | 365 | 365 | 365 | 365 |
| DB | 362 | 362 | 362 | 362 | 362 |
| QY | 425 | 425 | 425 | 425 | 425 |
| DB | 422 | 422 | 422 | 422 | 422 |
| QY | 485 | 485 | 485 | 485 | 485 |
| DB | 482 | 482 | 482 | 482 | 482 |
| QY | 545 | 545 | 545 | 545 | 545 |
| DB | 542 | 542 | 542 | 542 | 542 |
| QY | 605 | 605 | 605 | 605 | 605 |
| DB | 602 | 602 | 602 | 602 | 602 |
| QY | 665 | 665 | 665 | 665 | 665 |
| DB | 662 | 662 | 662 | 662 | 662 |

| | |
|------------|--|
| RESULT 9 | |
| AF363776 | |
| LOCUS | |
| DEFINITION | Heteractis crispa GFP-like chromoprotein mRNA, complete cds. |
| ACCESSION | AF363776 |
| VERSION | AF363776.1 GI:16660115 |
| KEYWORDS | |
| SOURCE | Heteractis crispa (leathery sea anemone) |
| ORGANISM | Heteractis crispa |
| | Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria; Stichodactylidae; Heteractis. |
| REFERENCE | 1 (bases 1 to 908) |
| AUTHORS | Gurskaya, N.G., Fradkov, A.F., Tersikh, A., Matz, M.V., Labas, Y.A., Martynov, V.I., Yanushevich, Y.G., Lukyanov, K.A. and Lukyanov, S.A. |
| TITLE | GFP-like chromoproteins as a source of far-red fluorescent proteins |
| JOURNAL | FEBS Lett. 507 (1), 16-20 (2001) |
| MEDLINE | 21538626 |
| PUBMED | 11682051 |

REFERENCE 2 (bases 1 to 908)
AUTHORS Gurskaya,N.G., Lukyanov,K.A., Labas,Y.A. and Lukyanov,S.A.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2001) Institute of Bioorganic Chemistry,
Miklukho-Maklaya 16/10, Moscow 117997, Russia
FEATURES
source
1..908
/organism="Heteractis crispa"
/mol_type="mRNA"
/db_xref="taxon:175771"
77..760
/note="hccp"
/codon_start=1
/product="GFP-like chromoprotein"
/protein_id="AAL27538.1"
/db_xref="GI:16660116"
/translation="MAGLLKESMRIMYMEGTVNGHVFKCEBGDGNPFTGTQSMRIH
VTEGAPLPFAFDILAPCEYGSRTFVHTAEIPDFKQSFEGTWTWTTTVEGGL
TAHQDTSLEGNCLYKVKVLGTFNFPADGPMKNKSGWEPCTEVVYPENGLVGRNV
ALKVGDRLICHLYTSYRSKAVRALTMPGFHTDIRLQMPKKKDEYFELYEASVAR
YSDLPEKAN"
ORIGIN
Query Match 63.6%; Score 436.6; DB 3; Length 908;
Best Local Similarity 77.5%; Pred. No. 4.5e-47;
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 5 TGAGCGGCCTGCTGAAGGAGAGATGTCGCATCAAGATGTACATGGAGGACCGTGAACG 64
Db 78 TGGCTGGTTTGTGAAGAAAGATGCGCATCAAGATGTACATGGAGGACCGTGAATG 137
QY 65 GCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA 124
Db 138 GCCATTATTCAAGTGTGAAGGAGGAGGAGCGCAACCCATTACAGGTACGCAGAGCA 197
QY 125 TGAGAAATCCACGTGACCGAGGGCGCCCGCTTCGCTTCGCTTCGACATCCTGGCCCCCT 184
Db 198 TGAGGATTGATGTCACCGAAGGGCTCCATTACCATTTGCTTCGACATTTGGCACCGT 257
QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTCTTCA 244
Db 258 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGCAGAGATTCCTGATTTCTTCA 317
QY 245 AGCAGAGCTTCCCCGAGGCTTACCTGGGAGAGAACCAACCACTACGAGGACGGCGCA 304
Db 318 AGCAGTCTTCCCTGAAGGCTTACTTGGGAAGAACCAACCACTATGAAGATGGAGCA 377
QY 305 TCCTGACCGCCACACAGGACACCGCTGGAGGGCAACTGCCTGATCTACAAGGTGAAG 364
Db 378 TTCCTACTGCTCATCAGGACACAAAGCCTGGAGGGCAACTGCCTTATATACAAGGTGAAG 437
QY 365 TGACCGGACCAACTTCCCGCGCAGCGCCCGTGTGATGAAGAACAAAGAGCGCGGCTGGG 424
Db 438 TCCTTGGTACCAATTTTCTGCTGATGGCCCGCTGATGAAGAACAAATCAGGAGGATGG 497
QY 425 AGCCAGACCGAGGTGTTACCCCGAGAACCGCGTGTGTGCGGCCGGAACGTGATGG 484
Db 498 AGCCATGCATGAGGTGTTATCCAGAGAAATGGTGTCTGTGTGGACGTAATGTGATGG 557
QY 485 CCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACAGCTACCGGAGCAAGA 544
Db 558 CCTTAAAGTCGGTGTATCGTCTGATCTGCCATCTCTATCTTACAGGTCCAAGA 617
QY 545 AGGCGGTGCGCCCTGACCATGCCCGGCTTCACCTTACCGACATCCGGTCCAGATGC 604
Db 618 AAGCAGTCCGTGCTTGAATGCCAGGATTTCAATTTACAGACATCCGCTTCAGATGC 677
QY 605 TGCGAAGAAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664
Db 678 CGAGGAAAAAGAAAGACGAGTACTTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG 737
QY 665 ACCTGCCCGAGAGGCCCAACTGA 687

Db 738 ATCTTCTCTGAAAAAGCAAAATTGA 760
RESULT 10
AX527888
LOCUS AX527888
DEFINITION Sequence 1 from Patent WO0230965.
ACCESSION AX527888
VERSION AX527888.1 GI:25172343
KEYWORDS
SOURCE
ORGANISM
Heteractis crispa (leathery sea anemone)
Heteractis crispa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
Stichodactylidae; Heteractis.
1
REFERENCE
AUTHORS Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent: WO 0230965-A 1 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source
1..910
/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"
ORIGIN
Query Match 63.3%; Score 435; DB 6; Length 910;
Best Local Similarity 77.3%; Pred. No. 7.3e-47;
Matches 528; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 5 TGAGCGGCCTGCTGAAGGAGAGATGTCGCATCAAGATGTACATGGAGGACCGTGAACG 64
Db 83 TGGCTGGTTTGTGAAGAAAGATGCGCATCAAGATGTACATGGAGGACCGTGAATG 142
QY 65 GCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA 124
Db 143 GCCATTATTCAAGTGTGAAGGAGGAGGAGGAGCGCAACCCATTACAGGTACGCAGAGCA 202
QY 125 TGAGAAATCCACGTGACCGAGGGCGCCCGCTTCGCTTCGCTTCGACATCCTGGCCCCCT 184
Db 203 TGAGGATTGATGTCACCGAAGGGCTCCATTACCATTTGCTTCGACATTTGGCACCGT 262
QY 185 GCTGCGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTCTTCA 244
Db 263 GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGCAGAGATTCCTGATTTCTTCA 322
QY 245 AGCAGAGCTTCCCCGAGGCTTACCTGGGAGAGAACCAACCACTACGAGGACGGCGCA 304
Db 323 AGCAGTCTTCCCTGAAGGCTTACTTGGGAAGAACCAACCACTATGAAGATGGAGCA 382
QY 305 TCCTGACCGCCACCAAGACACCGCTGGAGGGCAACTGCCTGATCTACAAGGTGAAGG 364
Db 383 TTCCTACTGCTCATCAGGACACAAAGCCTGGAGGGCAACTGCCTTATATACAAGGTGAAG 442
QY 365 TGACCGGACCAACTTCCCGCGCAGCGCCCGTGTGATGAAGAACAAAGAGCGCGGCTGGG 424
Db 443 TCCTTGGTACCAATTTTCTGCTGATGGCCCGCTGATGAAGAACAAATCAGGAGGATGG 502
QY 425 AGCCAGACCGAGGTGTTACCCCGAGAACCGCGTGTGTGCGGCCGGAACGTGATGG 484
Db 503 AGCCATGCATGAGGTGTTATCCAGAGAAATGGTGTCTGTGTGGACGTAATGTGATGG 562
QY 485 CCTGAAGGTGGCGACCGGACCTGATCTGCCACCACTACACAGCTACCGGAGCAAGA 544
Db 563 CCTTAAAGTCGGTGTATCGTCTTGTGATCTGCCATCTCTATCTTACAGGTCCAAGA 622
QY 545 AGGCGGTGCGCCCTGACCATGCCCGGCTTCACCTTACCGACATCCGGTCCAGATGC 604
Db 623 AAGCAGTCCGTGCTTGAATGCCAGGATTTCAATTTACAGACATCCGCTTCAGATGC 682
QY 605 TGCGAAGAAAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664
Db 683 CGAGGAAAAAGAAAGACGAGTACTTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG 742

Db 618 AACAGTCCGTCGCTTGACAAATGCCAGGATTTTCATTTTACAGACATCCGCTTCAGATGC 677
Qy 605 TCGGGAAGAAGAGGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664
Db 678 CGAGGAAACGAAAGACGAGTACTTTGAACCTGACGAAGCATCTGTGGCTAGGTACAGTG 737
Qy 665 ACCTGCCCGAGAGGCCCAACTGA 687
Db 738 ATCTTCCTGAAAAAGCAATGA 760

RESULT 13
AX527912
LOCUS AX527912
DEFINITION Sequence 25 from Patent WO0230965.
ACCESSION AX527912
VERSION AX527912.1 GI:251172357
KEYWORDS
SOURCE Heteractis crispa (leathery sea anemone)
ORGANISM Heteractis crispa
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
Stichodactylidae; Heteractis.

REFERENCE 1
AUTHORS Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.
TITLE Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL Patent: WO 0230965-A 25 18-APR-2002;
Clontech Laboratories Inc. (US)
FEATURES
source 1 .680
/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"

ORIGIN
Query Match 62.5%; Score 429.2; DB 6; Length 680;
Best Local Similarity 78.0%; Pred. No. 4.3e-46;
Matches 529; Conservative 0; Mismatches 148; Indels 1; Gaps 1;
Qy 10 GGCCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTCAACGGCCAC 69
Db 4 GGTGTGTTGAAAGAAAGTATGCGCATCAAGATGTACATGGAGGGCACCGTCAATGGCCAT 63
Qy 70 TACTTCAAGTGCAGGGCGGCGCCCTGCGCTTCGACATCCTGGCCCGCATGAGA 129
Db 64 TATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATTTCAGGTACCGACAGCATGAGG 123
Qy 130 ATCCACGTGACCGAGGGCGCCCTGCGCTTCGACATCCTGGCCCGCTGCTGC 189
Db 124 ATTCATGTACCGAAGGGGCTCCATTACCATTTGCTTCGACATTTTGGCACCGTGTGT 183
Qy 190 GAGTACGGCAGCAGGACCTTCGTGCACCAACACCGCGAGATCCCCGACTTCTTCAAGCAG 249
Db 184 GCGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCTCGATTCTTCAAGCAG 243
Qy 250 AGCTTCCCCGAGGGCTTCACCTGGGAGAGAACCAACCTACGAGGACGGCGGCATCCTG 309
Db 244 TCTTTCCTGAAGGCTTTACTTGGGAAGAACCACACCTATGAAGATGGAGGCATCTT 303
Qy 310 ACCGCCACACGAGACACCGCTGGAGGGCAACTGCTGATCTACAAGGTGAAGGTGCAC 369
Db 304 ACTGTTCATCAGGACACAGCTGGAGGGAACTGCTTATATACAAGGTGAAGTCTT 363
Qy 370 GGCACCAACTTCCCCCGCAGCGCCCGCTGATGAAGAACAAAGACGGCGGCTGGAGCCC 429
Db 364 GGTACCAATTTTCTGCTGATGGCCCGCTGATGAAGAA-AAATCAGGAGGATGGAGCCA 422
Qy 430 AGCACCAGGTGGTGTACCCCGAGAACGGCGTGTGTCGGCGCCGGAACGTATGGCCCTG 489
Db 423 AGCACTGAGGTGGTTTATCCAGAGAAATGGTGTCTGTGTGGACGTAATGTATGTCCTT 482
Qy 490 AAGTGGGCGACCGGCACCTGATCTGCCACCACCTACACAGCTACCGGAGCAAGAGGCC 549

Db 483 AAAGTCGGTGATCGTCGTTTGATCTGCCATCACTATACTTCTTACAGGTCCAAGAAAGCA 542
Qy 550 GTGCGCGCCCTGACCATGCCGGCTTCCACTTCCACCGACATCCGGCTCCAGATGCTGCGG 609
Db 543 GTCCGTGCTTGACAAATGCCAGGATTTTCATTTTACAGACATCCGCTTCAGATGCTGAGG 602
Qy 610 AAGAAGAAGGACGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCGACCTG 669
Db 603 AAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATCTGTGGCTAGGTACAGTGATCTT 662
Qy 670 CCCGAGAAGGCCCAACTGA 687
Db 663 CCTGAAAAAGCAATGA 680

RESULT 14
AY233273
LOCUS AY233273
DEFINITION Synthetic construct kindling fluorescent protein mRNA, complete cds.
ACCESSION AY233273
VERSION AY233273.1 GI:28629492
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 699)
AUTHORS Chudakov,D.M., Belousov,V.V., Zaraisky,A.G., Novoselov,V.V., Staroverov,D.B., Zorov,D.B., Lukyanov,S. and Lukyanov,K.A.
TITLE Kindling fluorescent proteins for precise in vivo photolabeling
JOURNAL Nat. Biotechnol. 21 (2), 191-194 (2003)
MEDLINE 22447893
PUBMED 12524551
REFERENCE 2 (bases 1 to 699)
AUTHORS Chudakov,D.M., Lukyanov,K.A. and Lukyanov,S.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2003) Institute of Bioorganic Chemistry, Mckluho-Maklaya, Moscow 117997, Russia
FEATURES
source Location/Qualifiers
1 .699
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
CDS
1 .699
/note="GFP-like protein; capable of photoconversion from nonfluorescent to red fluorescent form; KFP1; KFP-red"
/codon_start=1
/transl_table=11
/product="kindling fluorescent protein"
/protein_id="AAO43187.1"
/db_xref="GI:28629493"
/translation="MASLLTETMPFKTTEGTVNGHCFKICIGKEGNPFEGTQEMKIE VIEGGLPFAFHILSTSCMYGSKTFIKYVSGIPDYFKQSFPEGTWERTTTVEDGGFL TAHQDTSLDGLVYKVKILGNFPADGPMQNKVGRWEPGTEIYVEVDGVLRGQSLM ALKCPGGRHLTCHLHTTYRSKKPASALKMPGFHFEDHRIEIMEEVEKCKYQYEAAY GRYCDAAPSKLGHN"

ORIGIN
Query Match 61.8%; Score 424.8; DB 12; Length 699;
Best Local Similarity 79.3%; Pred. No. 1.6e-45;
Matches 530; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
Qy 5 TGAGCGGCTGCTGAAGGAGAGTATCGGCATCAAGATGTACATGGAGGGACCGTGAACG 64
Db 2 TGGCTCCCTGCTGACCGAGACCATGCCCCCTTCAAGACCACCATCGAGGGACCGTGAACG 61
Qy 65 GCACACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCCGCACCCAGAGCA 124
Db 62 GCCACTGCTTCAAGTGCATCGCAAGGGCGAGGGCAACCCCTTCGAGGGCACCCAGGAGA 121
Qy 125 TGAGAAATCCACGTGACCGAGGGCGCCCCCTGCCCTTCGCCCTTCGACATCTGCCCCCT 184
Db 122 TGAAGATCGAGGTGATCGAGGGCGCCCCCTGCCCTTCGCCCTTCACATCTCTCCACCT 181

This Page Blank (uspto)